CONCEPT AND DESIGN OF AN

AUDITORY LOCALIZATION CUE SYNTHESIZER

THESIS

RICHARD L. MCKINLEY

AFIT/GE/ENG/88D-29

DEPARTMENT OF THE AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

(1)

AFIT/GE/ENG/88D-29

CONCEPT AND DESIGN OF AN

AUDITORY LOCALIZATION CUE SYNTHESIZER

THESIS

RICHARD L. MCKINLEY

AFIT/GE/ENG/88D-29

Approved for Public Release; Distribution Unlimited

AFIT/GE/ENG/88D-29

CONCEPT AND DESIGN OF AN AUDITORY LOCALIZATION CUE SYNTHESIZER

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Electrical Engineering

RICHARD L. MCKINLEY, B.S.

December 1988

Preface

The purpose of this ~~study~~ thesis was to develop the concept and basic design for an auditory localization cue synthesizer. This technology has the potential for greatly reducing threat acquisition times in hostile ground-to-air missile scenarios by providing the pilot with a heads-up localizable auditory warning over his headset. This warning allows the pilot to quickly and naturally determine the location of the threat and take the necessary evasive actions. See p v

I wish to thank the many who assisted me in the development of this technology and writing this thesis. First, I would like to thank my wife Mary and my two children Betsy and Andy for their love, patience and understanding while working on this thesis. I would like to thank Drs Kabrisky, Nixon, Moore and Castor for their inspiration to excel and support of my research. Lt Mark Ericson deserves special recognition as my colleague at the lab. His tireless efforts in making the necessary electro-acoustic and human performance measurements made it possible to move the concept and design to reality. I would like to thank Mr David Ovenshire and Mr Ron Dallman for their work in the hardware and software area. Finally I would like to thank Hazel Watkins for the substantial efforts in typing this thesis.

RICHARD L. MCKINLEY

Table of Contents

List of Figures

AFIT/GE/ENG/88D-29

ABSTRACT

This thesis describes the concept and design of an auditory localization cue synthesizer. The pertinent literature was reviewed and used to form the basis of the concept a to generate localization cues over headphones utilizing real-time solid state processor. The synthesizer accepts a single monaural input and processes the signal separately for independent presentation to the left and right ears. The synthesizer uses a 3-space head tracking device to maintain a stable acoustic image when the listener moves his head. The design is complete to present localized stimuli in azimuth. A concept is described for generating stimuli in the three dimensional case for azimuth, elevation and distance. Details of the hardware and software design are in the appendices.

Laboratory methodology are described for deriving the necessary parameters of the synthesizer. Experimental data collected separately from this thesis demonstrate that the concept and design are viable for the azimuth case. Localization errors with the synthesizer are compared with free field errors obtained with 10 subjects. The results show that localization accuracy is essentially equal for the two conditions. Recommendations are presented for further research and development. (KR)

I. Introduction

Man can assimilate and operate on data optimally when the information is presented in a natural form. Acoustic information is normally presented binaurally in the natural world. These acoustic signals contain the cues that allow the listener, among other things, to discriminate the type of sound, the location of the sound source and the acoustic characteristics of the listening space. However, when listening with headphones, current audio systems present stereo acoustic signals which do not allow the listener to identify the location or estimate the distance of the sound. The ability to locate the source of a sound while wearing headphones would have a wide range of potential applications such as rapid target acquisition, multichannel conferencing (the cocktail party effect) and threat cueing.

The topic of this thesis is the concept and design of a real-time digital auditory localization cue synthesizer to generate the acoustic cues necessary to allow a listener to locate a sound source while listening with headphones.

Auditory localization is the ability to acoustically locate a single sound source relative to the listener, sometimes among several other sound sources, in azumith, elevation, and sometimes distance. The sound source is perceived to be outside the head and at some reasonable distance from the listener. These localized sensations are not normally perceived with signals generated by "stereo" listening with headphones.

Lateralization is the general sensation perceived by listeners of "stereo" sound with headphones. The lateralized signal is usually located at either the left ear or the right ear or somewhere in between but not outside the head of the listener.

Military applications of an auditory localization cue synthesis capability have been described as a necessary and integral part of the Project Forecast II, Super Cockpit and Virtual Man-Machine Interface project technologies. In the super cockpit, synthesized auditory localization is used in conjunction with helmet mounted displays. The auditory localization cue gives the pilot information that is outside his current field of view. This information can be anything from a threat warning from the radar warning receiver to an advisory to scan some of the displays not currently in the field of view. In addition, auditory localization has been projected to provide increased situational awareness by presenting auditory information and cues in a more natural and logical manner.

Commercial applications of an auditory localization cue synthesizer include "hi-fi" headphones since the synthesizer generates an out-of-head acoustic image, collision avoidance systems in commercial aircraft, navigation aids, video game applications, aids for the visually impaired and deep sea divers.

This thesis describes the concept and design of a device to synthesize localization cues over headphones. The goal of the device is to take a single audio input and process it independently for the left and right ears in such a manner that the resulting signal is localizable. The thesis details the rationale for the approach and the hardware and software necessary to realize this objective.

II. Background

Auditory localization has been researched for over 120 years. Throughout that period scientists have generally had the goal of understanding the mechanism of human auditory localization. Research has focused on the role of the pinna, interaural time delays, interaural intensity differences and head motion. Many different theories have been proposed of how humans localize sound. However, none-to-date satisfy all the well known experimental findings.

Fechner (12), in 1860 was one of the earliest researchers of mechanisms of human auditory localization. Batteau (2), in 1963 proposed a time delay theory of localization. He suggested that the pinna (the large cartilaginous portion of the external ear) introduced time delays to incoming sounds which allowed the auditory system to perform localization both monaurally and binaurally. Blauert (3), in 1969/1970 proposed that the pinna, head and ear canal caused angle of incidence dependent changes in the frequency spectrum of the sound source. This was generally called the theory of timbre differences. In 1974, Lambert (20) proposed the dynamic theory of sound source localization which is based on the effects of head movement on sound source azimuth and range. In binaural listening, Lambert proposed that interaural times were measured at the two ear locations by the auditory system to either map or calculate the location of the sound source. Kuhn (19) in 1977 reinforced these findings with his "Model for the Interaural Time Differences in the Azimuthal Plane". Kuhn used interaural time and interaural amplitude differences which showed that the KEMAR manikin gave data similar to that measured with human subjects. In addition Gatehouse (15) and Blauert (4) compiled books on "Localization of Sound: Theory and Applications" and "Spatial Hearing" respectively. Both books describe numerous investigations and how the various theories explain some but not all of the experimental findings.

The role of the pinna has been researched extensively. It is one of the major factors in the ability of humans to localize sound. It is the source of the frequency dependent interaural intensity differences. Batteau (1), in 1967 described the role of the pinna in human localization. He showed that it was physiologically possible that the time delays of 10 to 100 microseconds encoded by the pinna could be decoded by a simple neural net of excitation and inhibition. The role of the pinna in auditory localization was also described by Freedman (14) in 1968 who found that subjects with fixed head position were able to correctly localize sounds only when listening with either their own or artificial pinnae. The subjects had to move their heads to correctly localize sounds without pinna cues.

Shaw (31) over his lifetime has probably done the most extensive work on understanding the effects of the pinna on localization. He has performed detailed analysis of the external ear and the effects of the small anatomical features of the external ear on the transfer function of the pinna. In the future it may be possible to expand on Shaw's work and develop a computer model that would accurately predict pinna transfer functions from the geometry of the individual pinna. Shaw (32) also investigated the overall transform from free-space to the eardrum. This was reported in a paper published in 1974 which was a compilation of 12 studies. Wright (38) reported in 1974 that the pinna introduced time delays and that delays as small as 20 microseconds were perceptible by listeners. In 1975, Searle (29) proposed that differences between the two pinnae were used to localize sounds. If this is correct, it would indicate that left and right pinnae need to be measured and modeled independently in a localization cue synthesizer. Mehrgardt (21) in 1977 measured transfer functions of the external ear from 200 to 15000 Hz in both the horizontal and median planes. Moritmoto (23) published a paper in 1982 pointing out the importance of using the subject's own transfer functions in obtaining accurate localization. In 1984, Musicant (24) published a paper expounding on the fact that pinnae-based spectral cues were responsible for resolving front-back ambiguity in localization.

Clearly pinnae based cues play an important role in auditory localization. Any device to generate synthetic localization cues must accurately model pinna effects on the incoming signal. Burkhard (5) in 1975 described an acoustic manikin which accurately simulated acoustic diffraction of the head and torso and included pinnae and an eardrum simulator. This manikin called KEMAR, the Knowles Electronic Manikin for Acoustic Research, has received extensive use in the years since it became available. In 1969, Dirks (8) compared pinna transfer functions of

KEMAR with those measured by Shaw and found small differences only at high frequencies.

Pinna cues are so convincing that Hebrank (16), Flannery (13) and Musicant (25) found that two ears were not necessary for localization. Performance was less efficient in the monaural case but was enhanced when the subjects had an apriori knowledge of the spectrum of the sound source. In 1982, Colburn (6) working with subjects that were hearing impaired found that most subjects could localize within about 20 degrees using only one ear and, like Hebrank, found that apriori knowledge of the spectrum of the signal was required for optimum performance. Clearly, the pinna transfer functions are known to the listener much like a spatial map of an antenna pattern. Once the spectrum of the sound has been determined the human can use this map to determine the location of the sound source.

Interaural time delays have been described by a number of researchers as critical parameters for localization. Weiner (37) in his 1947 paper "On the diffraction of a progressive sound wave by the human head" found that interaural time delays alone were generally sufficient for localization in azimuth. Deatherage (7) in 1959, published a paper examining the trading relationship between interaural time delay and interaural intensity when localizing clicks and found that the relationship between the differences in intensity and time is not linear. Batteau (2), Blauert (3), Kuhn (19), and Doll (9) all found the interaural time delays ranging from 0 to 800 microseconds to be important in localization. Durlack (10) in 1986 published a paper in which the interaural time delays were increased giving the subject an illusion of listening with a head that was much larger than normal. This tended to give the listener the ability to more accurately determine the azimuth of a sound source.

A more natural method of increasing the accuracy of localization is by using head movement. Mills (22) in 1958 described the minimum audible angle for localization accuracy as being something on the order of 1 degree. Perrott (27) in 1981 showed that this 1 degree minimum audible angle held for moving sound sources up to 120 degrees per second. At an angular velocity of 240 degrees per second performance was degraded and the minimum audible angle increased.

The effects of head movement on auditory localization were described by Wallach (36) in 1940. His experiment showed that accurate head movements play an important role in localization. Thurlow (35) in 1967 showed that head movements reduced front-back reversals and supported Wallach's 1940 findings. Pollack (28) in his 1967 paper theorized that head movement was used to increase localization accuracy by moving the area of maximum sensitivity to the area of interest. Thurlow (34) supported this same finding in a 1967 paper. Lambert's (20) 1974 dynamic theory of localization embraced head motion as a critical parameter. In 1982, Shelton (33) described the role of vision and head motion in auditory localization. The major component in the effect was visually fixating on the apparent location of the sound source. Doll (9) in 1986 found, using a simulation of localization, that interaural time delays and head motion were the two critical parameters.

Throughout these scientific efforts, it is consistently apparent that the three critical parameters are, interaural time delay, frequency dependent interaural intensity and the dynamics of head motion. No one researcher has put all three parameters together to attempt the design of a localization cue synthesizer.

III. Concept

The concept for an auditory localization cue synthesizer is to generate over headphones in real-time the acoustic signals at the ears necessary for the listener to perceive the location of a sound source in space. The synthesizer is capable of processing a full range of acoustic signals and of maintaining an accurate and stable acoustic image during head movements of the listener. The location of the synthesized images presented to the listener is under the control of a host processor. Head position and movement are determined by a commercially available head position tracking system. The total integrated auditory localization cue synthesizer system provides synthesized cues in azimuth (horizontal), elevation (vertical) and distance (from the listener) for a complete three dimensional spatial localization environment.

Implementation of the concept utilizes the "brute force" method which consists of actual measurements of the acoustic transfer functions at the two ears for individual points in space located across the full ranges of azimuth and elevation as well as at selected distances. In very simple terms, these transfer functions which correspond to specific spatial locations are processed and stored in the auditory localization cue synthesizer. The stored processed signals are presented at the earphones under control of a host processor. and provide the listener with an image that appears to originate at the spatial location from which the signal was originally recorded

The total auditory localization cue synthesizer system is a laboratory demonstration breadboard system comprised of the auditory localization cue synthesizer itself, a head tracking system, a host processor and binaural headphones. The hardware and software required for these components and their interfaces are near the state-of-the-art in terms of processing speed and capacity. The basic configuration is shown in Figure 3-1. The host processor sends the auditory localization cue synthesizer the location or angle of the desired synthesized sound image. The system has three operating modes, azimuth only, azimuth and elevation and azimuth, elevation and distance. The values of the parameters of the desired locations are transferred by the host processor to the synthesizer over either an RS-232 or IEEE-488 bus. The RS-232 interface is adequate for azimuth only while the IEEE-488 bus interface with its higher data rate is required for the azimuth and elevation and azimuth, elevation and distance operations. A standard audio impedance of 600 ohms is provided by the synthesizer on both the input and output which is capable of handing ±10 volts. Head position information is provided by a commercially available Polhemus 3-space headtracker. This device measures head position in 6 degrees of freedom, x, y, and z position and roll, pitch, and yaw. The headtracker provides data output at a 54 Hz rate over a 16 bit parallel interface or at a 30 Hz rate over a RS-232C interface.

CONFIGURATION OF SYNTHESIZER SYSTEM

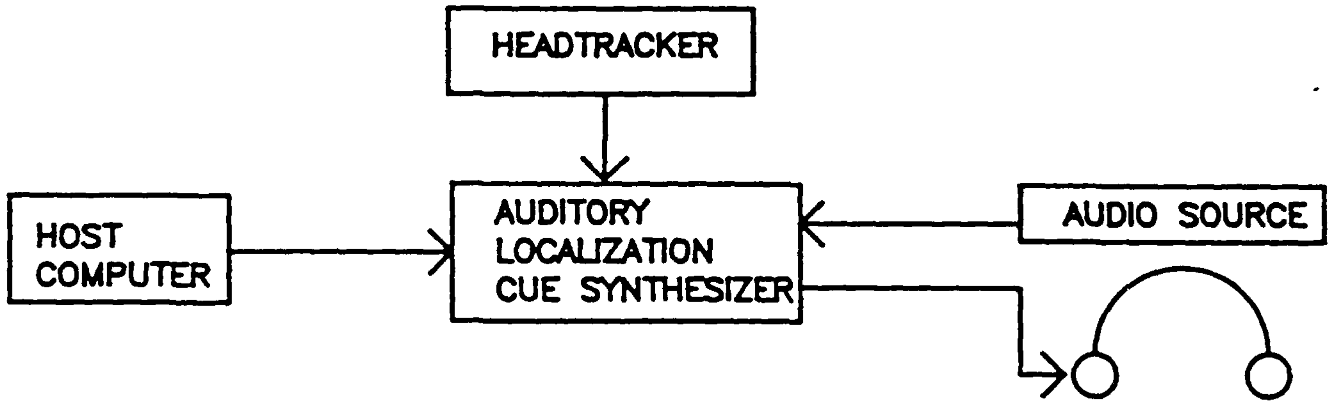


FIGURE 3-1

The auditory localization cue synthesizer was designed with the following goals.

1. Real-time operation
2. Minimum audible angle 1 degree (azimuth)
3. 10 kHz audio bandwidth
4. RS-232 interface to host processor
5. 16 bit parallel interface to headtracker.

The auditory localization cue synthesizer laboratory demonstration breadboard system is designed using currently available parts and is fabricated using wire wrap technology. The auditory localization cue synthesizer is divided into 3 functional subsystems.

1. Analog interface board
2. Digital interface board
3. Synthesizer processor board

This functional partitioning of the system allows for easy modification/upgrade of any one of the system functions without affecting the others. Figure 3-2 shows the internal interfacing between the various boards and the external interfacing.

The analog interface board has three external audio interfaces. The analog "in" port allows the desired location of the audio signal to be input to the system. The processed localizable signals are output independently for the left and right headphone channels. The internal analog interface board interfaces are three parallel interfaces of 16 bits each with necessary control and clock lines, one for the a/d

SYNTHESIZER INTERFACING

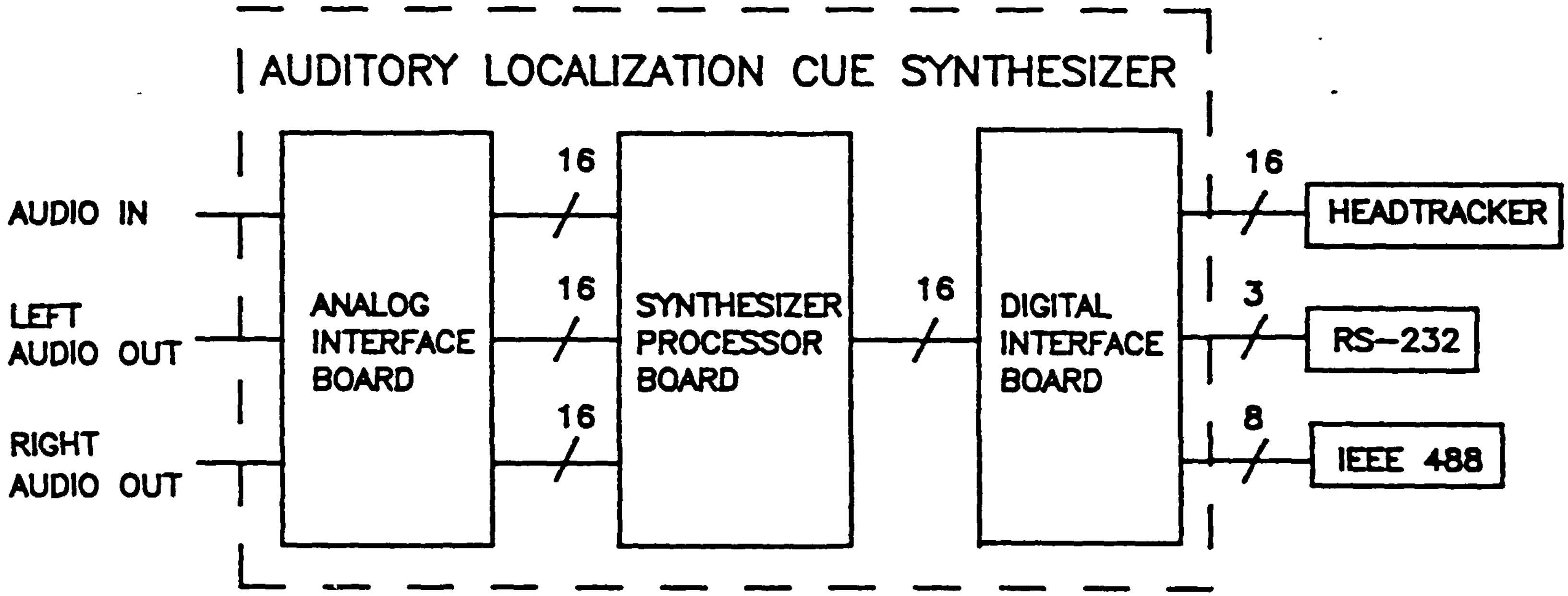


FIGURE 3-2

convertor and two for the d/a convertors that will be described in detail in Appendix A.

The digital interface board has two external interfaces, one 16 bit parallel interface for the headtracker and an RS-232C interface for the host processor. The internal interface is a single 16 bit parallel interface between the digital interface board and synthesizer processor board. The digital interface board contains a single interface processor to handle the maintenance of the headtracker interface and respond to commands from the host processor.

The synthesizer processor board is the core of the auditory localization cue synthesizer. The synthesizer processor board has two special purpose digital signal processors. Each digital signal processor provides angle information for the left and right ears. A second pair of digital signal processors provides distance information for the left and right ears. Once the algorithms described in Appendix B have been optimized it may be possible to reduce the hardware to a single digital signal processor for each ear to accomplish both the angle and distance processing.

Initially the signal processing operation has been bounded by synthesizing only the azimuth cues. Elevation and distance cues will be developed as logical progressions. The hardware is designed as much as possible to allow these developments to be simple software changes.

The three critical parameters in auditory localization are interaural time delay, interaural intensity cues (frequency dependent) and head motion. These cues are synthesized at one degree increments for the azimuth condition for relative angles of 0-359 degrees. This insures that each of the independent synthesized cues are separated by no more than the minimum audible angle. The interaural time delays were measured within the minimum 10 microsecond resolution capability of humans. The motion of the head has been measured at rates as high as 1000 degrees per second (11), however the Polhemus 3-Space headtracker which is the current state-of-the-art device outputs data at a maximum rate of 54 Hz. Subjects can move their heads almost 60 degrees per second without incurring a noticeable time lag. This is much more than the minimum audible angle of one degree. Head motion is tracked by updating the sound field at a 54 Hz rate. The time lag that occurs when the listener is rapidly moving the head is quickly resolved. Faster headtracking device are expected to be available in the near future. This headtracking method provides 360 independent cues for each ear for every angle in azimuth without interpolation.

The combination of azimuth and elevation synthesis requires a additional concept if the auditory localization cue synthesizer system is to operate with the same hardware for both the azimuth only, and azimuth and elevation conditions due to the limitations of memory space for the digital signal processor. The problem is how to provide one degree of resolution over a complete sphere while using about 360 points. The concept of an auditory fovea as shown in Figure 3-3 will provide the desired sensitivity without exceeding the memory space for

AUDITORY FOVEA

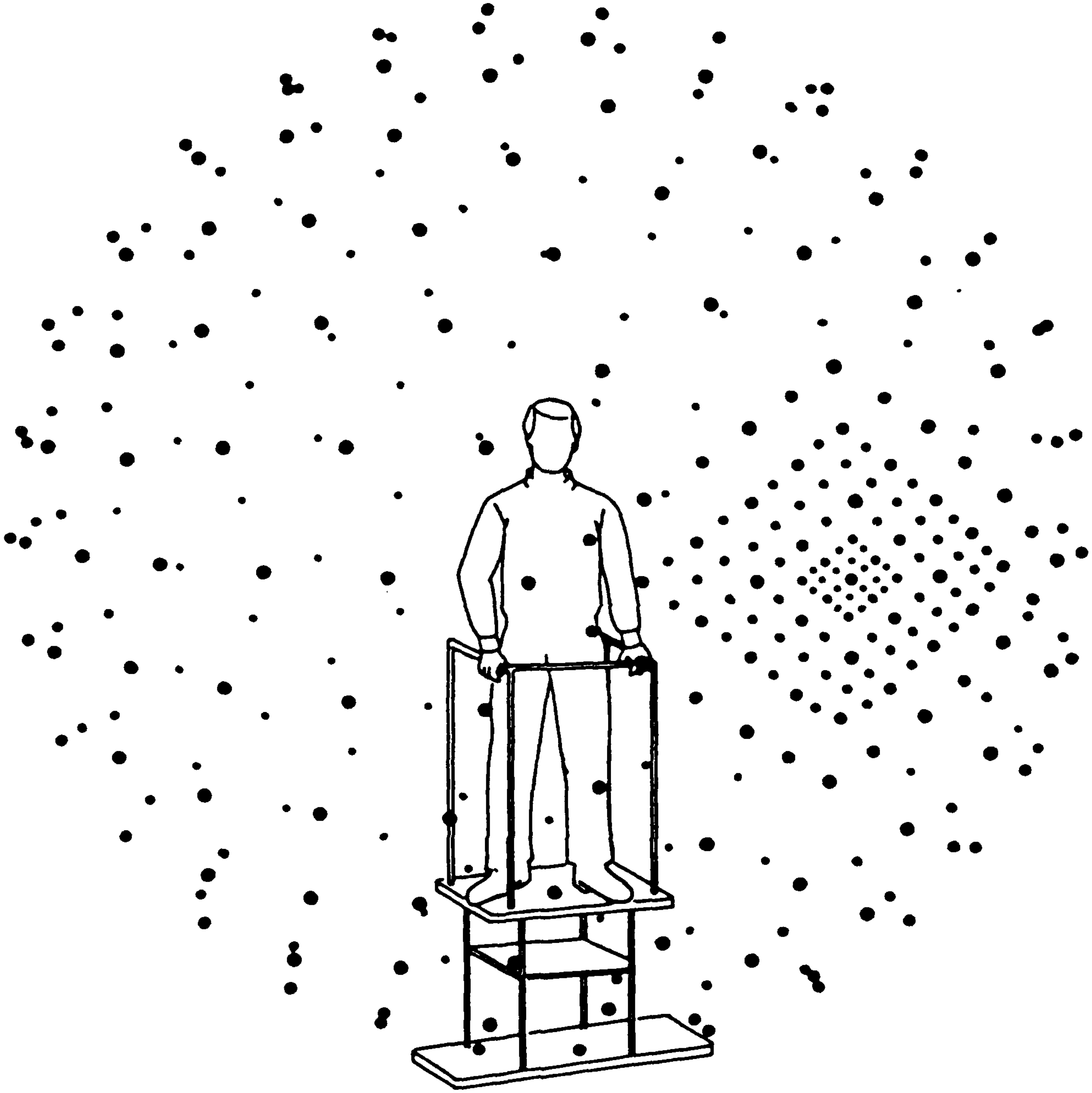


FIGURE 3-3

the digital signal processor. The auditory fovea is an area of increased resolution sensitivity directly in front of the observer ranging from 7.5 degrees to 1 degree. A maximum of 15 degrees is found everywhere else. When the maximum 15 degree spacing is used over the surface of a geodesic sphere, approximately 272 point sources are provided. The generation of the high resolution auditory fovea uses approximately 89 additional points for a total of 361 sources. The observer has high resolution anywhere his head is pointing with 15 degree resolution everywhere else when coupled with the headtracker.

Distance cues will be accomplished initially in a separate processor. The three parameters mapped into distance will be: attenuation of intensity, i.e. the inverse square law, low pass filtering effect of air and reverberation or multipath. The distance processor will synthesize the five different distances of very near, near, median, far, and very far. The synthesis of distance will operate on signals independent of the angle synthesis.

Care must be given to preserving as much signal quality as possible for these concepts to be successful. Signal quality is related to bandwidth, signal-to-noise ratio, distortion, and linearity. The design goals for this concept of an auditory localization cue synthesizer are as follows: bandwidth - 10 kHz, SNR - greater than 70 dB, distortion - less than 1 % and linearity - deviation less than 1 %.

In summary, the concept of the auditory localization cue synthesizer is based on a brute force time varying digital synthesis of the sound field of an external source at the entrance of the external ear canal. The three synthesis parameters are interaural time delay, interaural intensity, and head movement. The auditory localization cue synthesizer is based on currently available parts and is fabricated using wire wrap techniques. The azimuth only synthesis is realized with concepts and design presented for the elevation and distance cues.

IV. Approach

The auditory localization cue synthesizer was designed to provide auditory cues to persons wearing headphones allowing any audio signal to be localized under control of a host processor. The approach is direct and often referred to as "the Brute Force" approach. This procedure accurately describes each local point in space by a unique acoustic signal. In simple terms, the modifications of acoustic stimuli from points in space to the entrance of the ear canal are accurately modeled by the synthesizer. These modifications are measured as transfer functions which are resident in the synthesizer such that each one can be correlated with the signal to be localized, such that the processed signal over headphones is perceived to have originated from a point in space. In operation, the system is commanded to generate a localized signal in space, that point is compared with the current head position, a relative difference is derived and that difference is used to generate the desired signal with the transfer function and time delay associated with the relative angle and the output display over headphones.

The auditory localization cue synthesizer is essentially a real-time, time varying model of external physical acoustic signals as by the head, torso, and pinna. The software that runs on the auditory localization cue synthesizer is the model of these physical phenomena. The acoustic signals which provide the localization cues contain acoustic information as influenced by the head, torso and pinna of the observer. These acoustic factors must be accurately modeled for the observer to perceive the signal as being localized and out-of-head.

The measurements required to quantify the acoustic localization factors of a human subject are very time consuming. The measurements involving extensive mechanical positioning for each test point on a single set of pinna would require approximately one month for the collection of azimuth data. For this reason, an acoustically accurate manikin, KEMAR, (5), (8), was used as the model for measurements of the acoustic signals used in the synthetic cues in the auditory localization cue synthesizer. KEMAR's large orange (90th percentile) pinnae and 50th percentile head and torso are the bases for the generation of the synthetic cues in this auditory localization cue synthesizer.

A subject's ability to localize a sound is strongly influenced by the acoustic environment. A highly reverberant or diffuse sound field is one in which it is very difficult to localize because there is no apparent location to the source, i.e., the sound field is the same from every direction. Conversely, a highly absorptive or free field is one in which it is easy to localize. An anechoic chamber is a model of a free field, is highly absorptive, is free of significant reflections and has very low background noise. The anechoic chamber at AAMRL/BB was used to make the necessary acoustic measurements. This anechoic chamber has a noise floor that is 10 dB below the minimum audible field and any reflections by the walls are attenuated by at least 70 dB.

The pinna cues and time delay data are modeled using software that runs on a special purpose digital signal processor. The interaural time delay is modeled using a software controlled digital delay line. This interaural time delay measured in the AAMRL laboratory (11) ranges from approximately 0 to 750 microseconds. The interaural intensity differences for each point due to the pinnae are modeled using FIR filters, one for each ear. One FIR filter is used for each degree. There will be one FIR filter for each of the modeled points on the sphere for the azimuth and elevation case. The distance cue will be modeled using fractional multiplication for intensity, a low pass FIR filter to model the frequency specific attenuation of air and a 1.6 second long tapped delay line to model specific acoustic reverberation characteristics.

Analog to digital and digital to analog conversion are accomplished using commercially available 16 bit PCM convertors. The output of the digital to analog is amplified to directly drive a regular stereo headset. The digital interface board uses a digital signal processor as the controller interfacing with the synthesizer processor board, headtracker and host controller.

The complete synthesizer continuously provides audio signals in response to the direction commanded by the host processor. The desired direction can be either static or dynamic up to 54 desired directions per second. The system therefore generates a spatially stable acoustic image with or without head movement or a moving acoustic image with or without head movement.

V. Laboratory Measurements

The published literature on pinna cues, head related transfer functions and interaural time delay does not contain sufficient information to allow an auditory localization cue synthesizer to be designed and developed. Therefore it was necessary to design measurement procedures in the laboratory to describe the interaural time delay and head transfer function parameters. These parameters were measured in the Biodynamics and Bioengineering Division of the Armstrong Aerospace Medical Research Laboratory. The actual execution of the measurements (11) is not part of this thesis. This chapter describes the experimental equipment and test methodology used for measuring interaural time delay and head related transfer functions used in the synthesizer.

The KEMAR acoustic manikin (5) was used as the human model for both sets of measurements. KEMAR was positioned at the center of a 20 ft X 20 ft X 20 ft anechoic chamber, which has free field conditions down to a frequency of 63 Hz. KEMAR was mounted on a rotary stand calibrated in 15 minutes of arc over a 360 degree range. Each of the two ears was fitted with a free field microphone matched in both magnitude and phase. The microphones inputs were amplified by a matched pair of microphone preamplifiers. Both interaural time delay and head related transfer functions were measured at each degree from 0 to 359.

Interaural time delays were measured using the equipment set-up shown in Figure 5-1. A single loudspeaker was used to present the stimuli. The stimulus was one cycle of a 90% duty cycle 1 kHz triangle wave. This stimulus was selected because it delivered the maximum amount of information about the interaural time for the equipment used. The outputs of KEMAR's two ears were recorded by a two channel 12 bit digitizing oscilloscope with a sampling rate of up to 4 MHz. The 1 MHz sampling rate was used for data collection to get the required 10 microsecond resolution and both channels were digitized simultaneously. The time differences between the two ears in microseconds were averaged over ten stimuli. After each measurement KEMAR was rotated one degree and the next measurement made until all data were collected. The interaural time delay data collected and the interaural sample delay value used in the auditory localization cue synthesizer for the azimuth only case are shown in Appendix C. The time delay data were divided by 25 microseconds and rounded to the nearest integer to arrive at the sample delay value. In a similar fashion, time delays can be measured for the azimuth and elevation case. However KEMAR must be elevated and translated for positions off the equator of the sphere being modeled for these measurements.

Head related transfer functions were measured with the experimental set-up in Figure 5-2 using a single loudspeaker to present the stimuli. Several methods of determining the transfer functions in both magnitude and phase were evaluated. One method used a two channel spectrum analyzer

INTERAURAL TIME DELAY MEASUREMENTS IN AZIMUTH

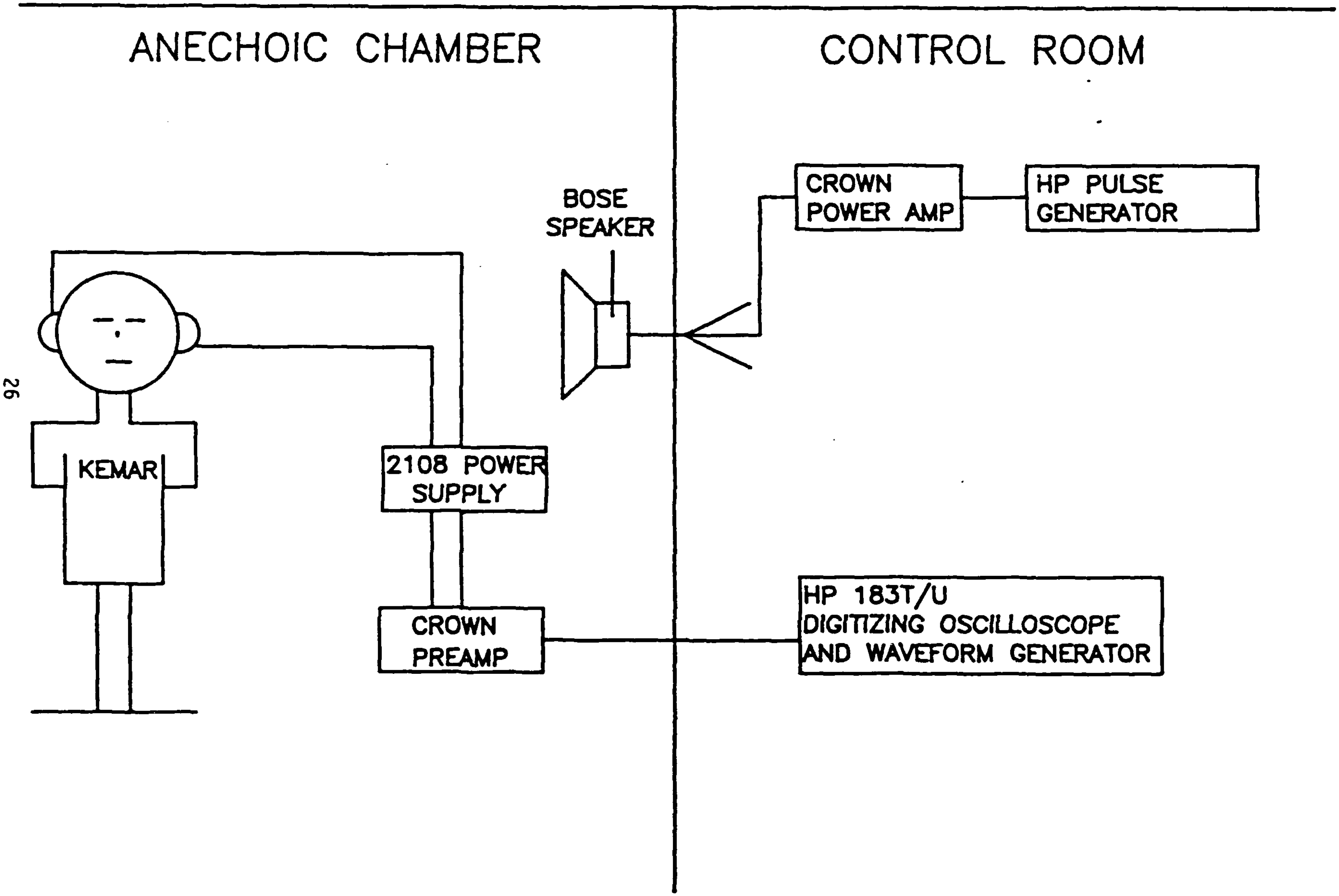


FIGURE 5-1

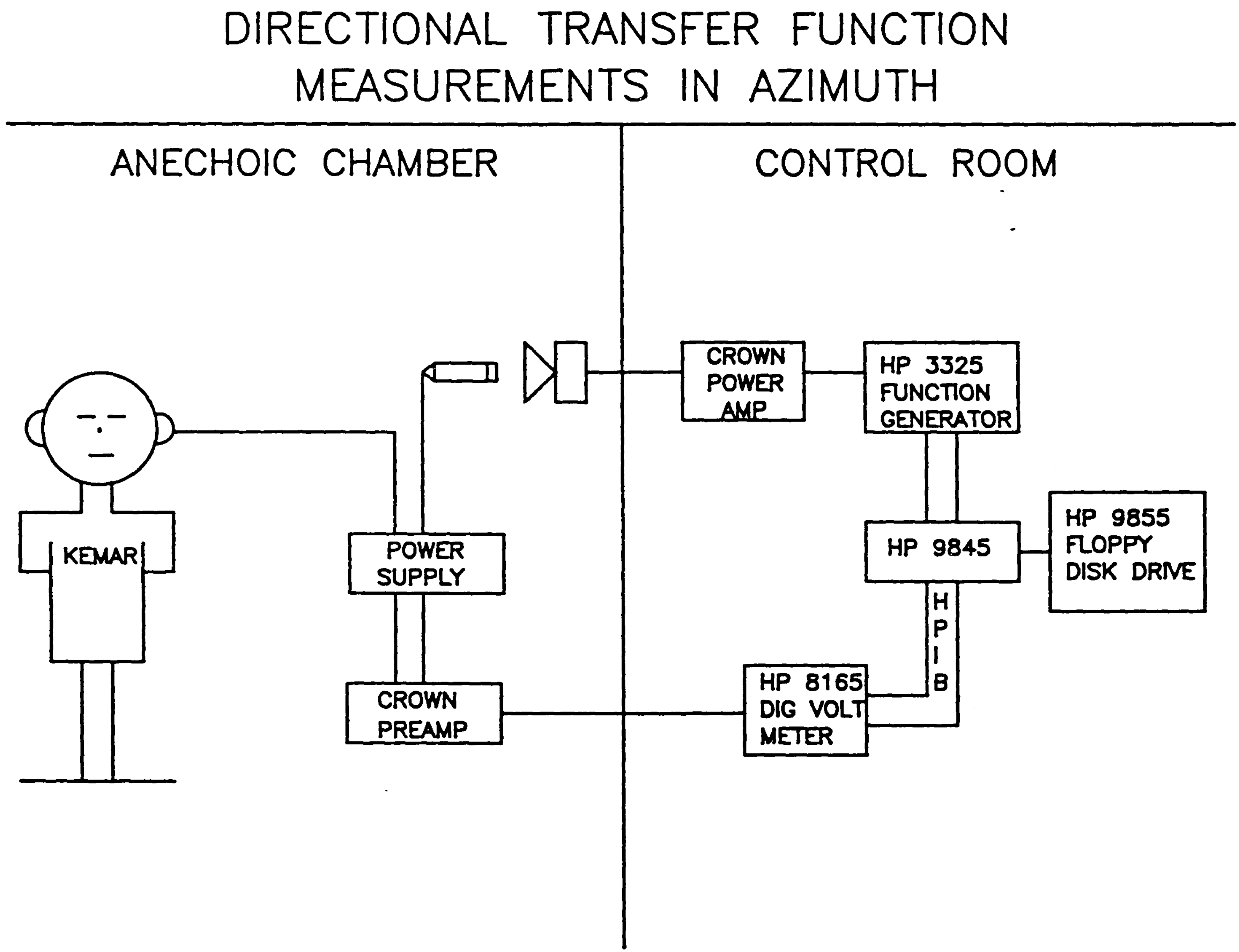

DIRECTIONAL TRANSFER FUNCTION
MEASUREMENTS IN AZIMUTH
ANECHOIC CHAMBER
CONTROL ROOM
KEMAR
POWER SUPPLY
CROWN PREAMP
CROWN POWER AMP
HP 3325 FUNCTION GENERATOR
HP 9845
HP 9855 FLOPPY DISK DRIVE
HPIB
HP 8165 DIG VOLT METER

FIGURE 5-2

with a logarithmic swept sine stimulus. It was very time consuming and once the correction was made in the phase response for the time of the propagation delay for sound in air, there was very little information left in the phase response. Another method used and found satisfactory was a single channel analysis technique with a logarithmic swept sine input. The logarithmic data collection more accurately models the response of the human auditory system to frequency than does linear data collection and the interaural time delay model effectively takes care of the phase response. Head related transfer function data were collected from 100 Hz to 20 kHz in 1/12 octave steps, for every degree from 0 to 360 for the azimuth only case. Appendix D shows representative samples of the measured head related transfer functions. These transfer functions were modeled in the auditory localization cue synthesizer using a 179 tap FIR (26) filter designed using the Kaiser window (18)(30). An alpha=1 parameter was used in the design of the 720 filters to maximize their accuracy in both frequency and amplitude. Each filter used quantized 16 bit coefficients for maximum signal to noise ratio in the fixed point processing of the TMS320C25. The head related transfer functions for the azimuth and elevation can be measured using the same elevation and translation method described for interaural time delay measurement.

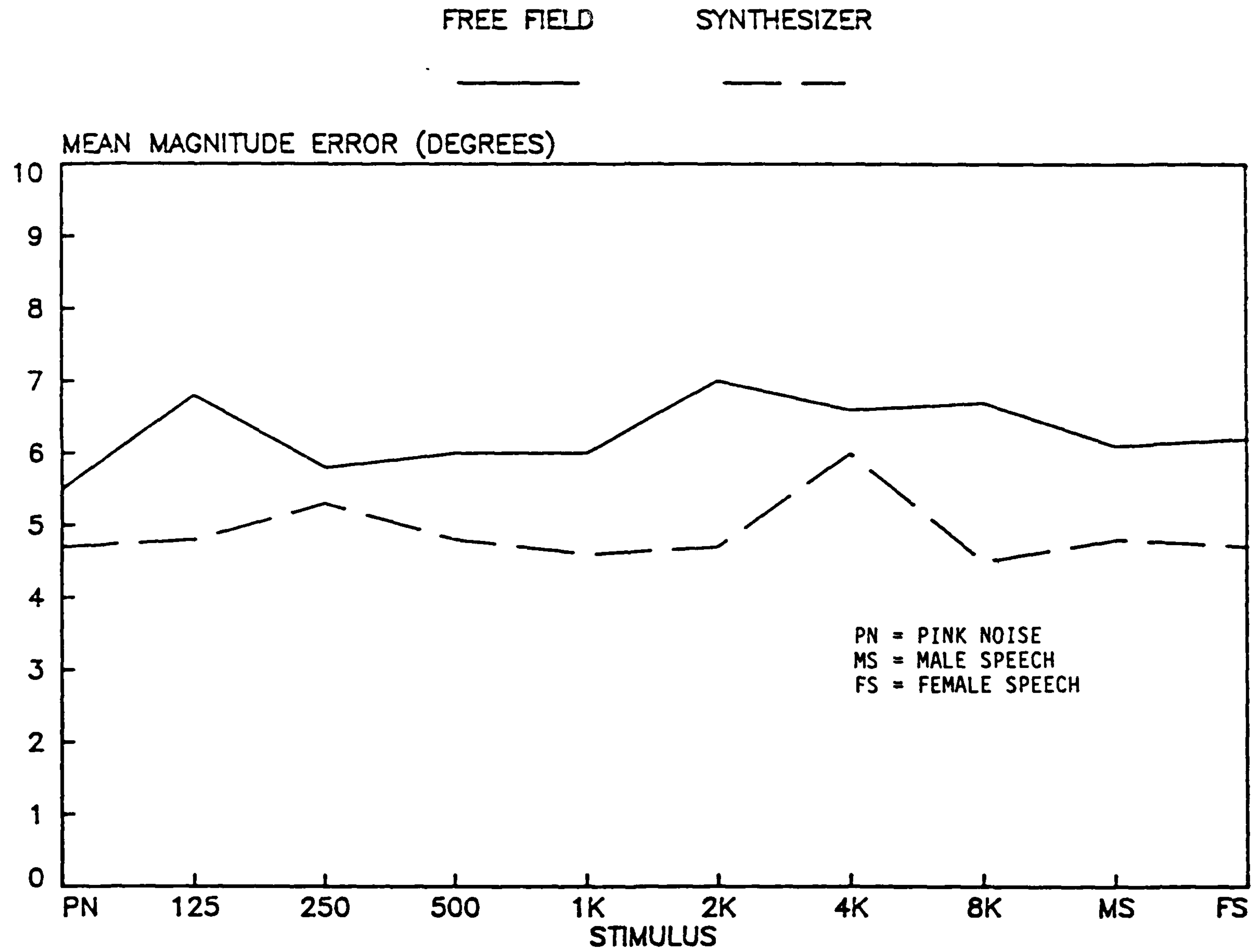
The procedures described provided sufficient data to implement an auditory localization cue synthesizer. Interaural time delays ranged from 0 to almost 750 microseconds. Head related transfer functions were measured in magnitude only using a logarithmic swept sine technique in 1/12 octave band steps from 100 Hz to 20 kHz. The overall gain of the transfer functions ranged over 10 dB. These electro-acoustic data form the parameters of the model which the localization cue synthesizer implements. Other models based on other heads, pinnae, and torsos can likewise form the basis for localization cue synthesizer.

VI. SUMMARY AND RECOMMENDATIONS

A concept and design of an auditory localization cue synthesizer have been presented. The concept for the localization synthesizer included implementing in real-time three parameters known to influence localization, interaural time delay, head related transfer functions or pinna cues and correction of the first two cues for head motion. These parameters are synthesized using a high speed digital signal processing system. The interaural time delay is implemented with a software delay line and the head related transfer functions are implemented using 360 different 179 tap FIR digital filters. The synthesizer is capable of responding to commands and imparting the requested localization cue information to regular stereo headphones. The complete localization synthesis system consists of the auditory localization cue synthesizer, a head tracker, a host computer or controller and a pair of headphones.

The concept and design presented herein have been fabricated by AAMRL and SRL for the azimuth only case outside the scope of this thesis. Initial human localization performance measurements conducted separately from this thesis (11) indicate that the synthesizer localization performance, Figure 6-1, is comparable with human performance in a free-field. These findings demonstrate that the concept and design presented in this thesis are viable for at least the azimuth only case. Further work is required before the azimuth, elevation and distance cases can be evaluated. AAMRL efforts are

HUMAN AUDITORY LOCALIZATION PERFORMANCE
MEAN MAGNITUDE ERROR BY STIMULUS



BASED ON DATA OF 10 SUBJECTS
FIGURE 6-1

currently underway to fabricate the azimuth and elevation synthesizer to include distance cues.

The auditory localization cue synthesizer concept has many applications. Cockpit applications are of primary interest to the Air Force. Auditory 3-D displays could be generated to give threat warnings with directional and distance information over the pilot's headset removing some of the overload from the visual system. In addition, intelligibility of monitoring multiple radio transmissions could be increased (17) by spatially separating the different channels. Other applications include air-air collision avoidance system cues over headphones, situational awareness aids by employing a 3-D auditory display, a spatial disorientation recovery aid by providing an auditory orientation cue, airborne, undersea and ground based virtual display systems and 3-D auditory displays for air traffic controllers. Commercial applications include home "hi-fi," audio-video games, and entertainment.

Research applications of the auditory localization cue synthesizer are extensive. The auditory localization cue synthesizer can generate cues over headphones that are impossible to generate in a free-field. Examples include generation pinna cues without any interaural time delay or time delay cues without pinna information. The apparent size of the head could be altered by software modification. Idealized pinnae transforms to possibly enhance localization performance could be investigated. Experiments conducted using this technology should give new insight into the mechanism of human auditory localization. Possibly this information could lead to a unified theory of human auditory localization.

Recommendations for future work include completing the azimuth and elevation case and adding the distance cues to complete the verification of the concept. Additional research is required to determine the optimum audio bandwidth for localization in azimuth only and in azimuth and elevation. The effects of listening with pinnae different from those of the listener is a major question. If humans can satisfactorily adapt to other pinnae, then individual transfer functions would not be required for practical applications. Work is needed in the area of multiple source and dynamic source localization. Finally, better understanding of the mechanism of human auditory localization should allow the development of better localization cue synthesizers. This thesis is just a beginning. The scientific responsibility is to continue to push back the frontier.

Appendix A

Hardware Design

The auditory localization cue synthesizer was designed in functional partitions to facilitate debugging and future upgrades without a complete redesign. The three primary functional partitions are (1) analog interface board, (2) digital interface board and (3) synthesizer processor board.

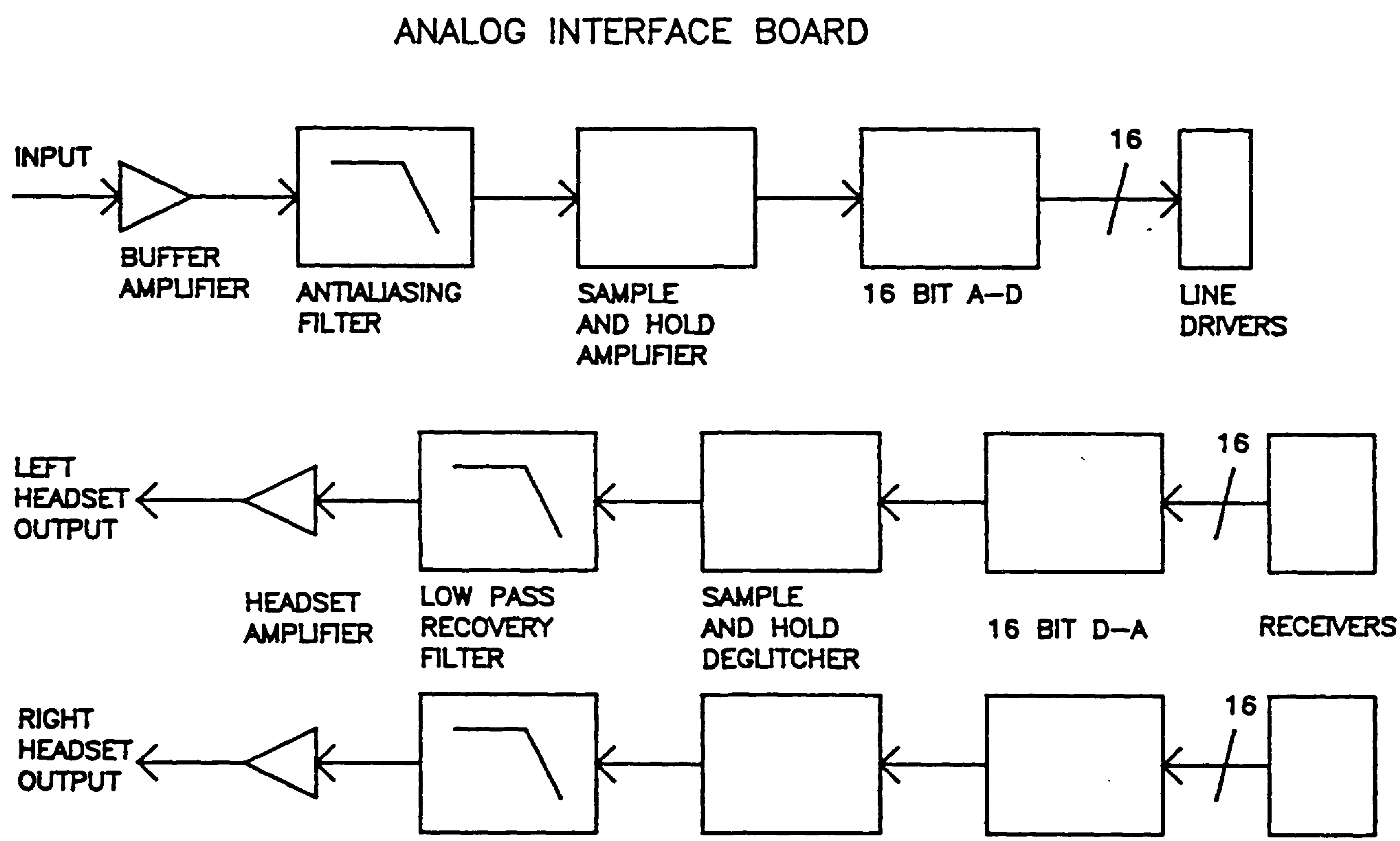
The system was designed to run at the full capacity of each of the major components to allow maximum flexibility in manipulation of the software based algorithms. The system has resident on ROM all software to execute the localization algorithms and relies on the host processor only for commands for the desired location of the source.

Analog Interface Board

The analog interface board is designed to perform the a-d and d-a conversions for the auditory localization cue synthesizer and its basic functional diagram is seen in Figure A-1. It supports a maximum 10 kHz audio bandwidth and is subdivided into an input section, and left and right output sections.

Input

The input section consists of an input buffer amplifier, antialiasing filter, sample and hold amplifier, and a-d. The entire input system is designed to support input signals up to ±10 V while maintaining a noise floor less than 153 microvolts. (a) The input



ANALOG INTERFACE BOARD
INPUT
BUFFER AMPLIFIER
ANTIALIASING FILTER
SAMPLE AND HOLD AMPLIFIER
16 BIT A–D
16
LINE DRIVERS
LEFT HEADSET OUTPUT
HEADSET AMPLIFIER
LOW PASS RECOVERY FILTER
SAMPLE AND HOLD DEGLITCHER
16 BIT D–A
16
RECEIVERS
RIGHT HEADSET OUTPUT
16

FIGURE A-1

buffer amplifier protects the antialiasing filter, sample and hold amplifier, and a-d from over voltage and matches impedance for the antialiasing filter to 600 ohms. The OP-27 was selected for the analog interface board for its high gain-bandwidth product and low noise floor. (b) The antialiasing filter attenuates any signal to one-half the least significant bit of the a-d at frequencies of one-half the sampling rate and greater. The Frequency Devices antialiasing filter used attenuates the signal 85 dB in one octave and uses the Cauer elliptic design. The 3 dB corner frequency is set at 10 kHz. This design gives 104 dB of attenuation at 20 kHz and 85 dB attenuation above 20 kHz. The specified 16 bit a-d has a 96 dB range. In a worst case, 9 dB of aliasing would be present if a 40 kHz sampling rate is selected. This is an acceptable amount of worst case aliasing for a laboratory prototype system. (c) The Analog Devices sample and hold amplifier SHA-1144 used in the design has worst case specs good only to 14 bits. However at the time of the design it was the best sample and hold amplifier available in the speed required. The function of the sample and hold amplifier is to sample the signal in a very short time and then hold the signal stable while the successive approximation a-d converts the analog voltage to the 16 bit PCM representation. The a-d takes 17 microseconds to convert in the worst case. Therefore the sample and hold amplifier must hold the analog signal stable within one-half the least significant bit of the a-d during the 17 microsecond conversion. This calculates to a droop rate of 153 microvolts per 17 microseconds or less than 10 microvolts

per microsecond. The SHA 1144 takes 6 microseconds nominally and 8 microseconds maximum to acquire a signal. Therefore the total sampling and conversion time max is 25 microseconds, which gives a sampling rate of 40 kHz. Figure A-2 shows the relationship between the 10 kHz desired bandwidth and the 40 kHz sampling rate. As can be seen, with a potential broad band input, and realistic antialiasing filter the 40 kHz sampling rate is required to support the 10 kHz bandwidth. The noise of the SHA-1144 is 70 microvolts peak-peak which is less than the 153 microvolts of one-half the least significant bit of the a-d convertor.
(d) The 16 bit a-d convertor selected is the Burr-Brown PCM75. This a-d was specifically designed for audio applications. Its maximum conversion time is 17 microseconds and has less than 0.004% total harmonic distortion. It is a successive approximation a-d with 16 bit parallel outputs.

The a-d convert line and the sample and hold amplifier control line are under control of the angle processor to be described later. The purpose for this arrangement is to allow the angle processor software control of the sampling rate. This allows easy changes of system bandwidth by modifying the software and changing the antialiasing filter.

In summary, the input section of the analog interface board includes four basic components, an input buffer amplifier, an antialiasing filter, a sample and hold amplifier, and a 16 bit analog-to-digital convertor. This input section reduces the bandwidth

BANDWIDTH VS SAMPLING RATE

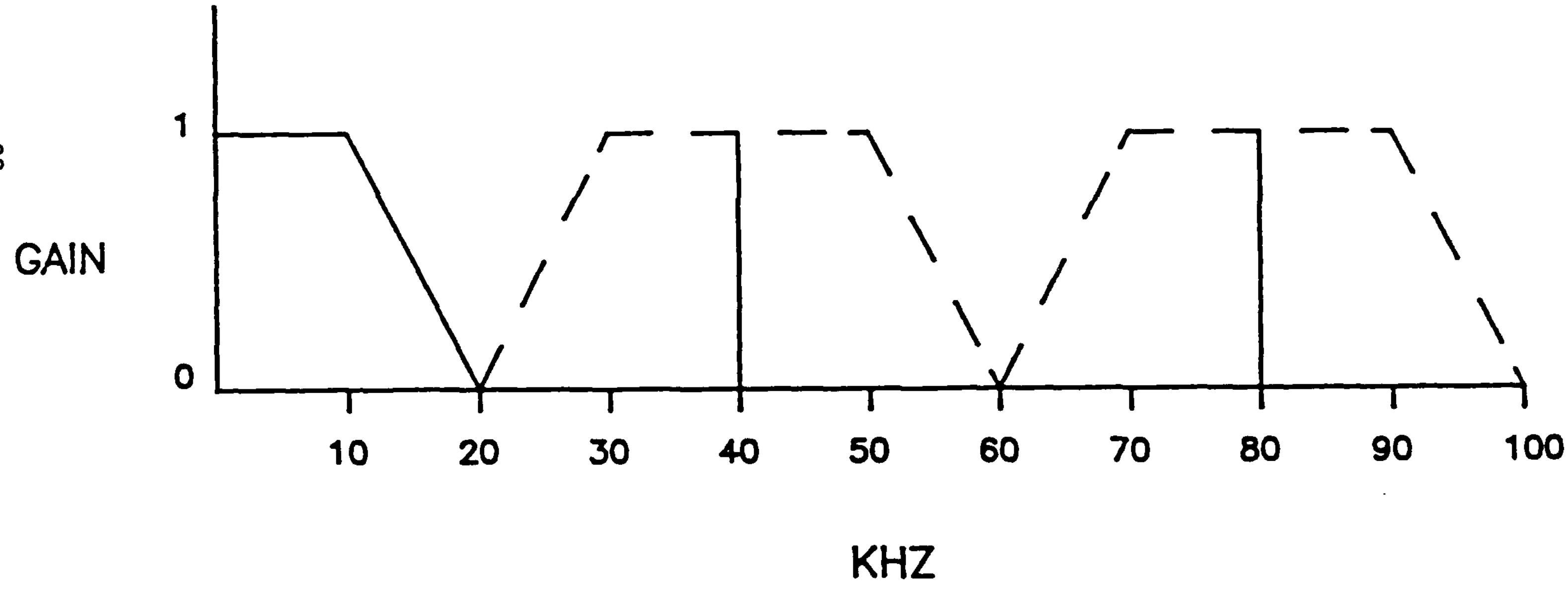


FIGURE A-2

of incoming signals to 10 kHz to prevent aliasing when sampled at 40 kHz. The total input section has an 85 dB signal-to-noise ratio.

The output section of the analog interface board contains an independent left and right channel. Each channel has a separate d-a, sample and hold amplifier, antialiasing filter, and headset amplifier. The output channel sampling rate is under software control up to a maximum of 40 kHz. The maximum voltage output is ±10 V into an 8 ohm load.

The output section operates on the same ±10 V analog signal as the input section while maintaining a noise floor less than 153 microvolts. (a) The PCM-53 16 bit d-a converter selected for the analog interface board was specially designed by Burr-Brown for audio applications. The settling time is 4 microseconds or less and the d-a has less than 0.002% total harmonic distortion with a full scale input. The d-a output signal is unstable during the 4 microsecond settling time. (b) An Analog Devices SHA-144 sample and hold amplifier is used to deglitch the output by holding the previous voltage while the new value settles, then the sample and hold amplifier switches to the sample mode. The 4 microseconds of the d-a and the 8 microseconds of the sample and hold amplifier give a worst case d-a conversion time of 12 microseconds. This is much less than the 25 microseconds required to support the 40 kHz sampling rate. The output conversion, like the previously described input conversion, is under software control. (c) The Frequency Devices lowpass recovery filter is used to smooth the stairstep output of the d-a within the original 10 kHz bandwidth. An 8 pole Butterworth filter with a 10 kHz corner frequency was selected for the lowpass recovery filter. (d) A National Semiconductor monolithic 5 watt audio amplifier drives the output of the lowpass recovery filter to a standard 8 ohm impedance or higher audio headset. In addition, the unamplified output of the lowpass recovery filter is available for separate component amplification.

The analog interface board is interfaced to the synthesizer processor board by ribbon cables. The three 16 bit a-d and d-a parallel interfaces communicate over the ribbon cables via line drivers and receivers. In addition, the ribbon cables carry the necessary clock and control lines.

The arrangement of components of the analog interface board is designed to insure signal fidelity up to 10 kHz within the ± 10 V range. It provides the analog interface for the synthesizer processor board. It supports software controlled sampling rates up to 40 kHz. The input and output a-d and d-a converters are 16 bits each. The input and output channels contain antialiasing and lowpass recovery filters, and sample and holds and deglichers.

Digital Interface Board

The digital interface board is the subsystem that is essentially the I/O processor for the auditory localization cue synthesizer. The digital interface board has a separate digital signal processor that acts as the I/O processor. This I/O processor is independent of the synthesizer processor board and implements the interface with the headtracker and host processor. The function of the processor is to accept desired angle commands from the host processor on an interrupt basis, accept real head angles from the headtracker on a periodic basis and provide the synthesizer processor board computed relative angle indexes on a periodic basis. The I/O processor insures that the relative angle indexes are available before being needed by the synthesizer processor board. The interface to the host processor is implemented in two ways, one an RS-232C interface and the second an IEEE-488 interface. The RS-232C interface is simple to implement, is a common interface, and is limited to 9.6 kbits per second. The IEEE-488 interface is somewhat more difficult to implement, is a less common interface than RS-232C, but offers speeds up to about 500 kbits per second. The RS-232C interface is satisfactory for stationary and slowly moving stimuli (less than 180 degrees per second). The IEEE-488 interface is required for high speed dynamic stimuli (greater than 180 degrees per second). The interface with the synthesizer processor board is implemented as two separate 16 bit parallel I/O ports with semaphores, one each for the left and right synthesizer processors.

The digital interface board is organized as shown in Figure A-3. The five I/O ports on the I/O processor consist of three 16 bit parallel interfaces, a serial interface and an 8 bit paralleled interface. The program is contained in ROM and uses off-chip scratch pad memory. For the azimuth only case the I/O processor is a TMS-32020 digital signal

DIGITAL INTERFACE BOARD

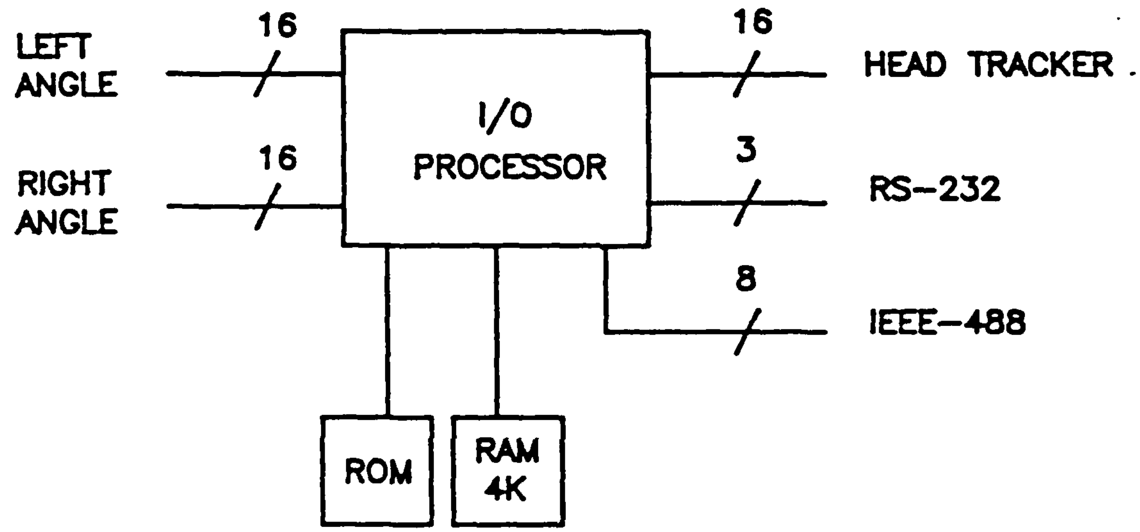


FIGURE A-3

processor. For the azimuth and elevation case, a TMS-320C25 is required, because of the order of magnitude greater computing requirements. The two chips are pin-for-pin compatible and only require a simple clock chip change with an appropriately designed board.

Synthesizer Processor Board

The analog interface board and the digital interface board provide an interface to the outside world for the synthesizer processor board. The synthesizer processor board contains the two digital signal processors for implementing the digital models of human auditory localization cues. Each of the digital signal processors functions independently and is independently interfaced with the analog interface board and digital interface board. Figure A-4 shows the functional layout of the synthesizer processor board. In order to obtain maximum fidelity of the synthesized localization cues, state of the art TMS320C25 digital signal processors were used. These digital signal processors have a nominal 100 nanosecond instruction cycle time. One of the most common operations in digital signal processing is multiply and accumulate. The TMS320C25 does this combination operation with memory fetches and register decrements in a single 100 nanosecond cycle. This allows long FIR filters to be implemented in real time. The TMS320C25 has on-chip ROM and RAM in addition to supporting off-chip ROM and RAM. The memory requirements of the auditory localization cue synthesizer dictate that the full address space of off-chip ROM and RAM be used. The program and data are stored in slow ROM and then booted to high

SYNTHESIZER PROCESSOR BOARD

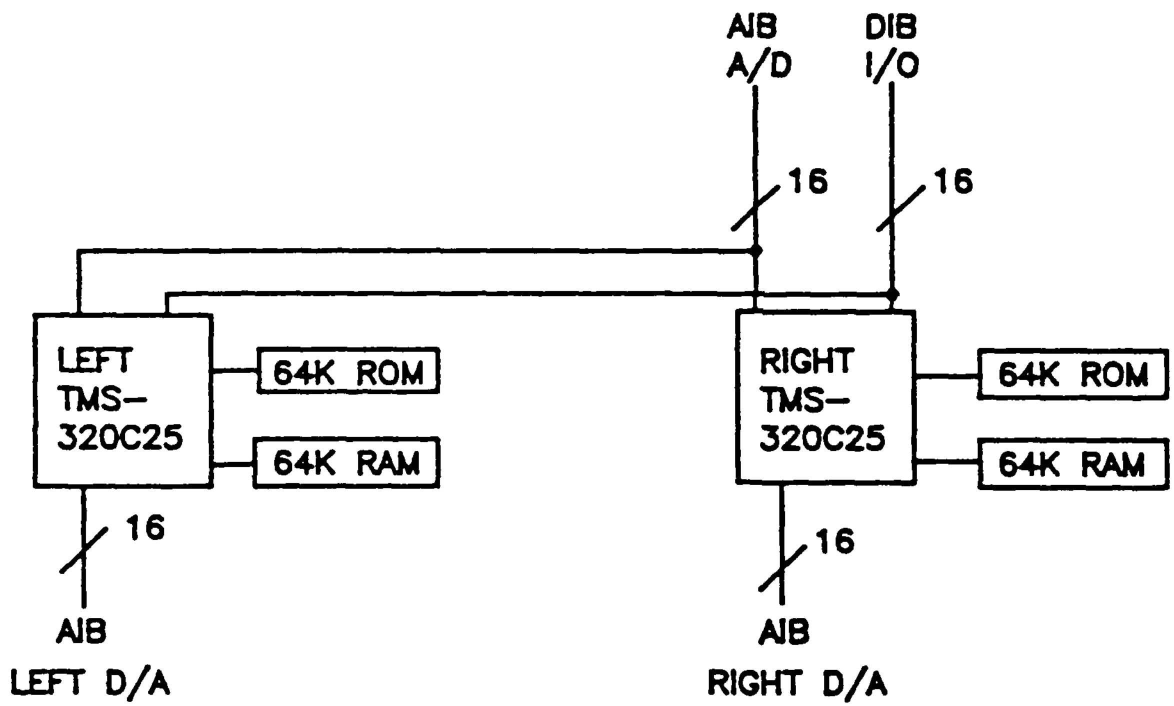


FIGURE A-4

speed (25 ns) RAM upon power-up or reset. This allows easy modification of software by using slow UV erasable ROM's while allowing full speed operation from the high speed RAM. The TMS320C25 has a modified Harvard architecture that has separate program and data memory spaces. During the transfer of the program from ROM to RAM a software operated switch on one of the I/O ports makes the RAM look like data memory to the processor. Once the boot is completed the switch makes the RAM look like program memory and execution from RAM begins. The two processors are running the same program with different data. If humans' two pinnae were exactly the same, a common memory could be shared instead of two separated banks of memory. The interfaces to the analog interface board and digital interface board are all 16 bit parallel interfaces with semaphores. The hardware configuration for the azimuth only and azimuth plus elevation conditions are the same on the synthesizer processor board.

In summary, the auditory localization cue synthesizer hardware is designed in blocks to match the three functional partitions. An analog interface board forms the interface to the outside world with a single audio input and independent left and right audio outputs. A digital interface board is the interface processor between the host, headtracker, and the synthesizer processor board. A synthesizer processor board is the heart of the auditory localization cue synthesizer and actually implements the synthesis algorithms which generate the localization cues.

Appendix B

Software Design

The auditory localization cue synthesizer software implements the localization cue models in real-time. All code is written in TMS-320 assembly language in order to optimize speed of execution. This chapter describes the software for the angle processors on the synthesizer processor board. The software for the I/O processor on the digital interface board was part of an earlier joint AAMRL - SRL Laboratory effort. Software for the distance processor will be a future effort.

The synthesizer processor board software implements 360 independent FIR filters that model the transfer function from free-space to the entrance to the ear canal. Each angle processor has an individual set of 360 sets of coefficients so that differences between left and right pinnae can be modeled. In addition, each processor implements in software a 30 sample (25 microseconds per sample) delay line that models the interaural time delays. Each set of coefficients for the 360 FIR filters has associated with it a number from 0 to 30 for the interaural time delay. In the azimuth only case the 360 points are used to implement models of each degree in azimuth. In the azimuth and elevation case, 272 points are used to implement a model of a sphere at maximum 15 degree spacing with the remainder of the points implementing an auditory fovea in the users' high localization resolution area. All

MEMORY MAP AFTER CNFD

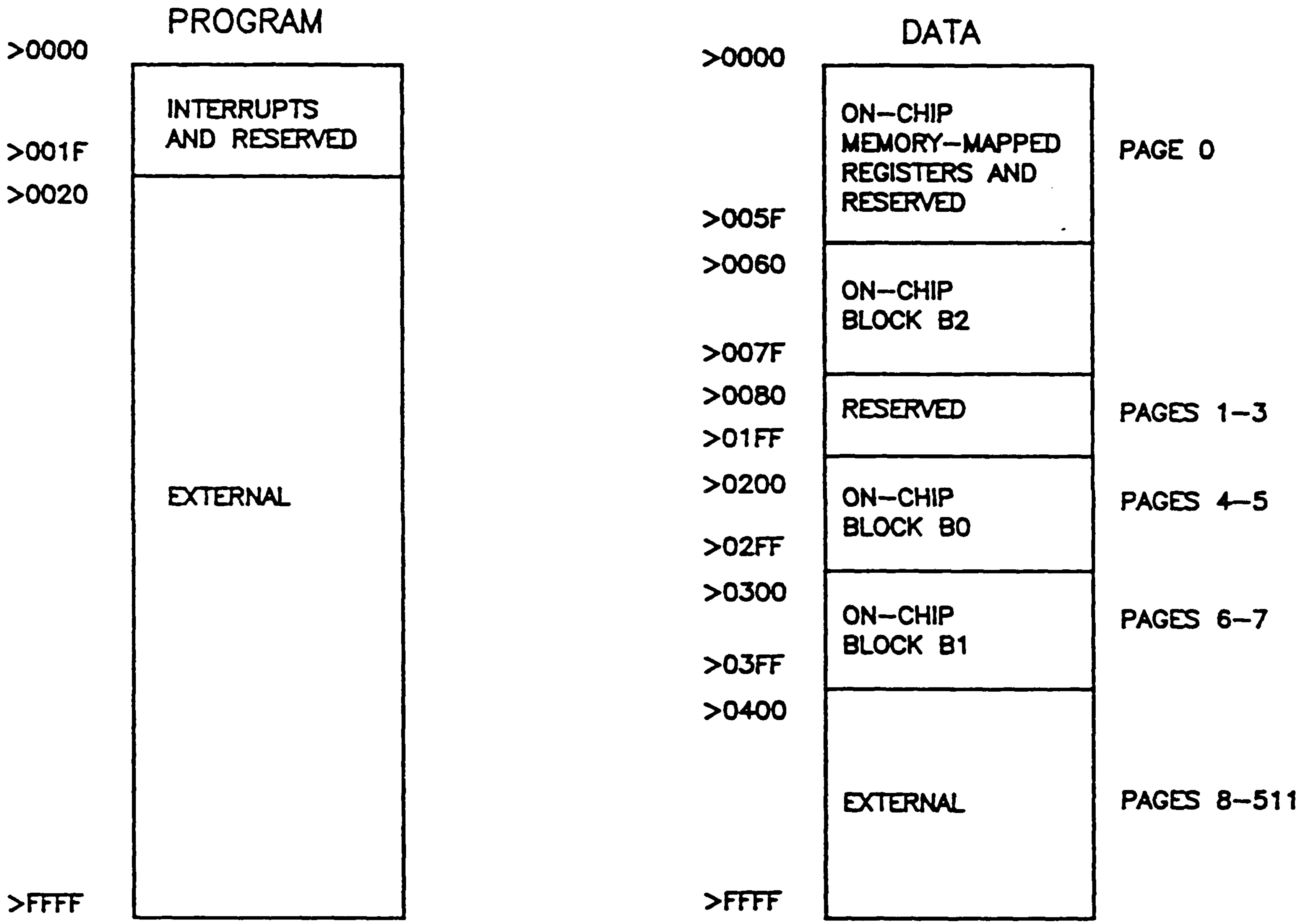


FIGURE B-1

MEMORY MAP AFTER CNFP

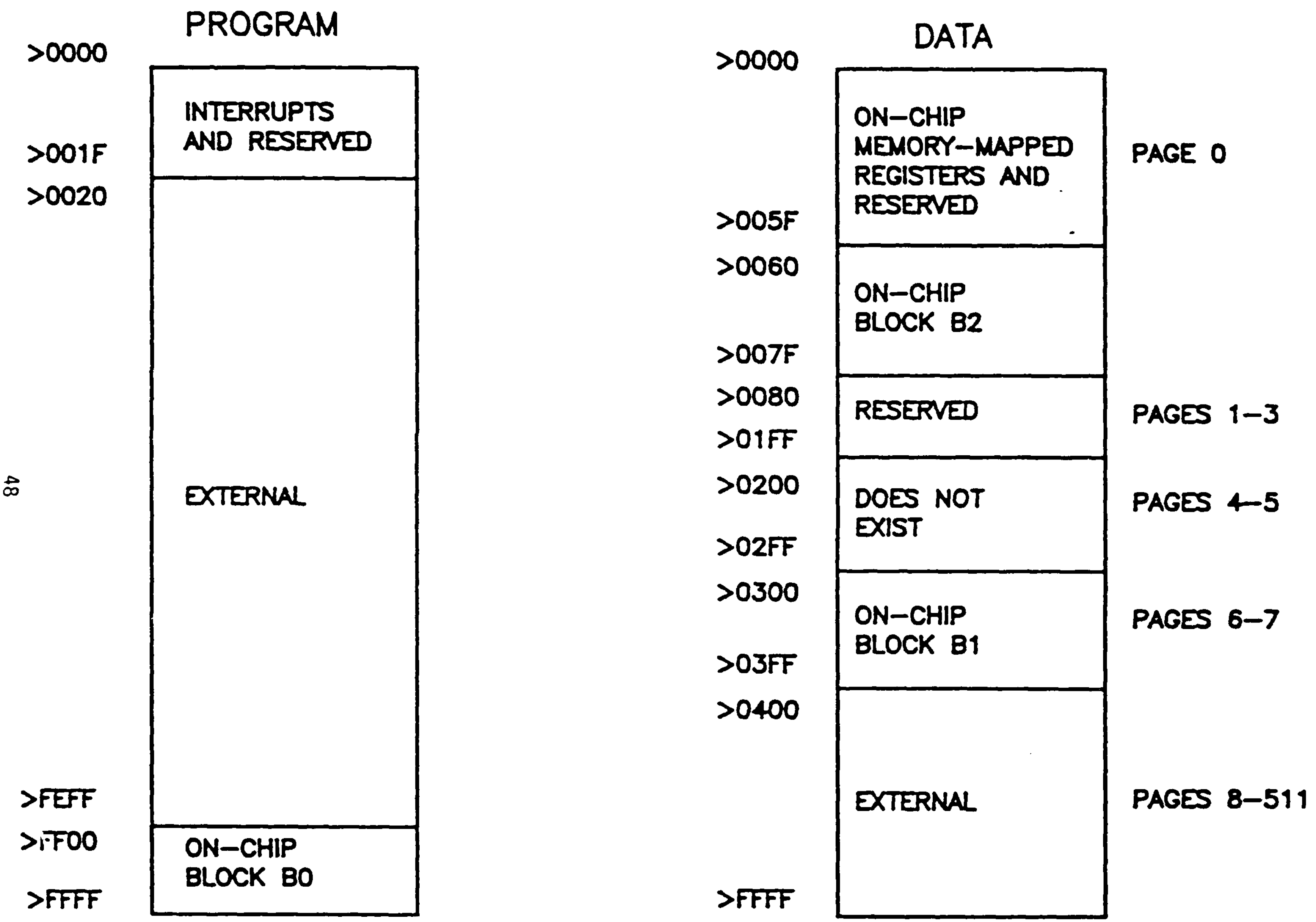


FIGURE B-2

FLOW CHART INITIALIZATION

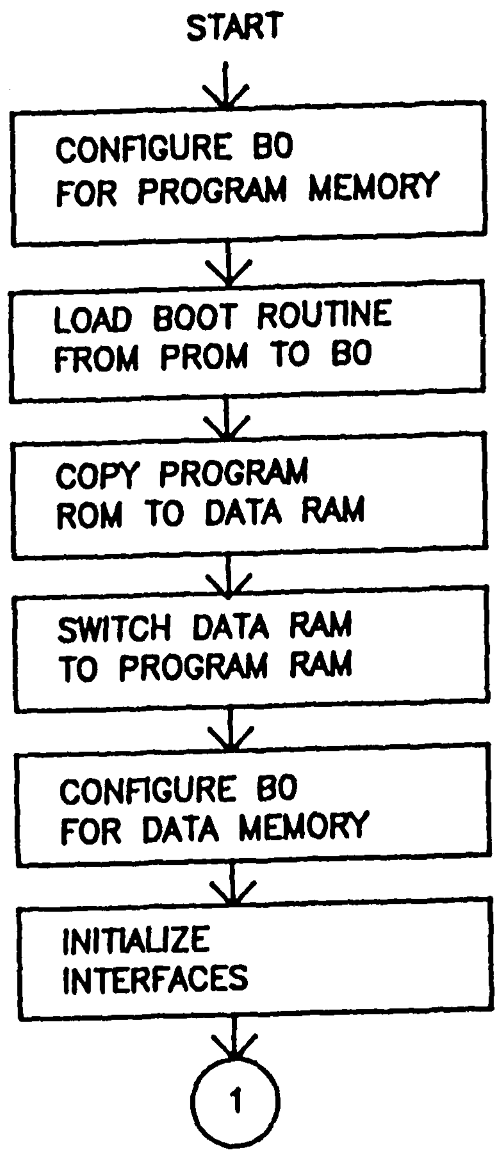


FIGURE B-3

FLOW CHART - LOCALIZATION CUE SYNTHESIS

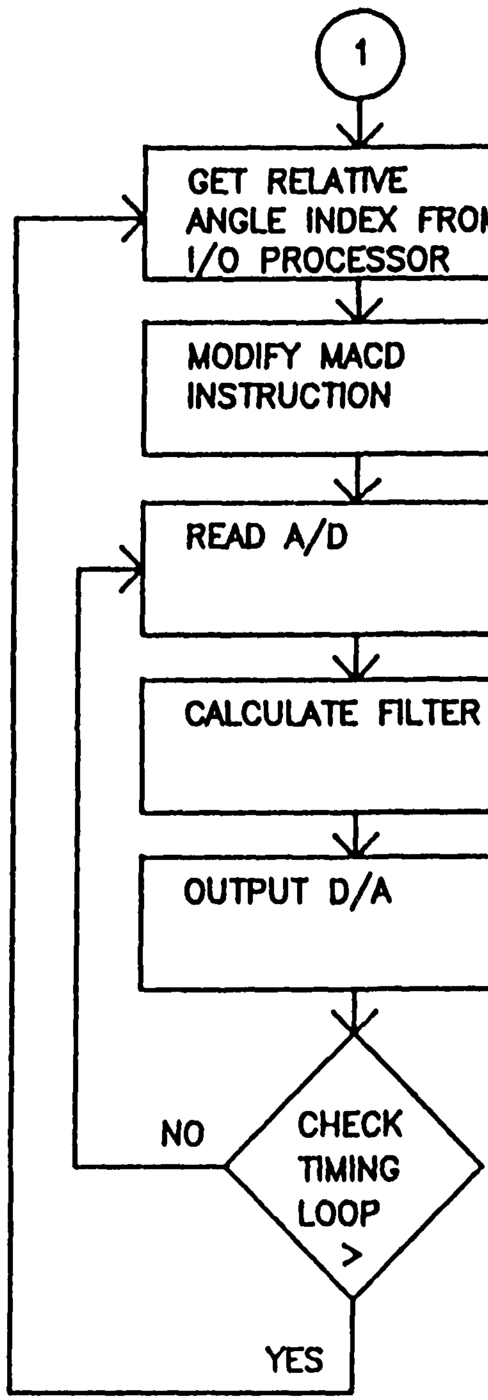


FIGURE B-4

software is resident in UV erasable ROM then downloaded to RAM upon system power up.

The TMS320C25 has 64 K words (128 K bytes) of program memory, some on-chip and some off-chip. It also has 64 K words (128 K bytes) of data memory, some on-chip and some off-chip. The allocation of a block of on-chip memory can be changed from program to data or data to program under software control. The CNFD instruction configures memory block B0 for data memory and the CNFP instruction configures B0 for program memory. The memory maps for the two configurations are shown in Figures B-1 and B-2. In addition, the auditory localization cue synthesizer hardware can output a 0 over an I/O port which configures the high speed RAM as program memory or an output of a 1 which configures the high speed RAM as data memory. This feature allows the entire data memory to be copied to program memory and allows the program memory to be modified during run time.

The angle processor software includes a boot routine, an initialization routine, a self-modifying instruction routine and a filter and delay routine. Figures B-3 and B-4 show the angle processor flow chart. The execution of the modules is as follows. The boot routine initially runs in program ROM and immediately copies itself to on-chip RAM which has been configured for program memory. The boot routine configures external RAM as data memory and copies the program, FIR filter coefficients, and delay values from the program ROM to data RAM. The last instruction of the boot routine switches data RAM to program RAM and program execution proceeds from RAM. A CNFD instruction is executed to configure on-chip RAM block B0 as data RAM. The on-chip RAM locations that will be used for the input data (179 locations) and the output delay line (30 locations) are zeroed.

Initialization is completed by enabling the overflow mode, disabling interrupts, and enabling the sign extension mode.

The self-modifying instruction routine is called from the filter and delay routine. The filter and delay routine begins by clearing the I/O processor interface then getting the next relative angle index from the I/O processor. This index is then multiplied by 180 (the number of FIR coefficients plus the delay value) and added to an address offset to get the absolute address of the first coefficient of the FIR filter for that angle. This address is then used to modify the pointer address portion of the MACD (multiply, accumulate and decrement pointers) instruction to use the proper set of coefficients in external program RAM. During the actual modification, the RAM is switched from program memory to data memory and the instruction modified. The RAM is then switched back to program memory with the pointer address portion of the MACD instruction modified.

The filter and delay routine continues and a value is then read from the a-d and 179 multiplies and accumulates form the product. During the multiply and accumulate, all values in the a-d table are moved one location in preparation for forming the next product. The products are stored in a 30 location output delay line. The value output is the one indexed by the delay value associated with the FIR filter calculated. The DMOV statement is used to move the data on location higher in memory each time a new product is added. NOP instructions are used to accurately time the loop to exactly 25 microseconds. A loop counter is checked, if the I/O processor is not ready with a new index then another a-d value is read. If the I/O processor has a new value it is read and a new filter calculated.

The auditory localization cue synthesizer angle processor software implements 360 FIR filters of 179 taps in real time. A 30 location delay line and the FIR filters are indexed by a single value between 0 and 359 transferred from the I/O processor to the angle processor. The software implements the total algorithm in less than 25 microseconds allowing real-time operation.

APPENDIX C

INTERAURAL TIME DELAY

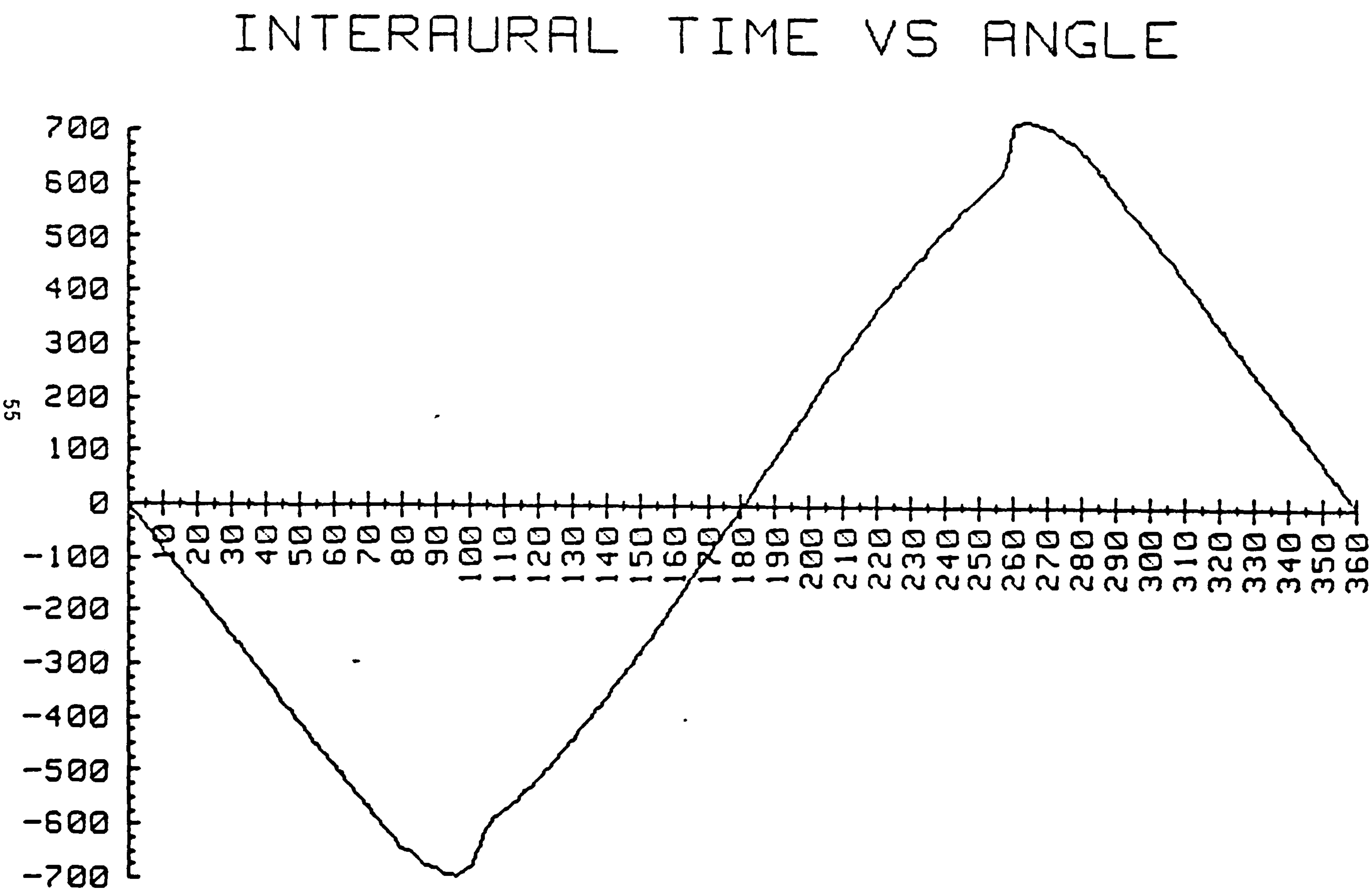

INTERAURAL TIME VS ANGLE
700
600
500
400
300
200
100
0
-100
-200
-300
-400
-500
-600
-700
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350
360

FIGURE C-1

Interaural Time Delay

ANGLE	INTERAURAL TIME DELAY (uSEC)	QUANTIFIED INTERAURAL TIME DELAY SAMPLES AT 40 KHz (25 uSEC PER SAMPLE)
1	-7	-0
2	-17	-1
3	-23	-1
4	-34	-1
5	-42	-2
6	-50	-2
7	-55	-2
8	-65	-3
9	-73	-3
10	-86	-3
11	-93	-4
12	-100	-4
13	-108	-4
14	-116	-5
15	-122	-5
16	-131	-5
17	-141	-6
18	-146	-6
19	-156	-6
20	-163	-7
21	-169	-7
22	-182	-7
23	-189	-8
24	-196	-8
25	-206	-8
26	-213	-9
27	-220	-9
28	-229	-9
29	-239	-10
30	-247	-10
31	-254	-10
32	-260	-10
33	-270	-11
34	-279	-11
35	-287	-11
36	-295	-12
37	-305	-12
38	-310	-12
39	-319	-13
40	-328	-13

41	-335	-13
42	-345	-14
43	-354	-14
44	-361	-14
45	-373	-15
46	-377	-15
47	-385	-15
48	-393	-16
49	-402	-16
50	-410	-16
51	-415	-17
52	-427	-17
53	-438	-18
54	-444	-18
55	-449	-18
56	-458	-18
57	-469	-19
58	-475	-19
59	-481	-19
60	-489	-20
61	-500	-20
62	-504	-21
63	-514	-21
64	-521	-21
65	-529	-21
66	-537	-22
67	-545	-22
68	-550	-22
69	-559	-23
70	-565	-23
71	-577	-23
72	-585	-24
73	-592	-24
74	-601	-24
75	-606	-25
76	-614	-25
77	-622	-25
78	-628	-25
79	-637	-26
80	-642	-26
81	-646	-26
82	-649	-26
83	-654	-26
84	-658	-27
85	-665	-27
86	-669	-27
87	-672	-27
88	-676	-27

89	-670	-27
90	-679	-27
91	-683	-28
92	-689	-28
93	-691	-28
94	-690	-28
95	-693	-28
96	-695	-28
97	-691	-28
98	-689	-27
99	-683	-27
100	-678	-27
101	-668	-26
102	-654	-25
103	-635	-25
104	-615	-24
105	-605	-24
106	-594	-24
107	-588	-23
108	-583	-23
109	-577	-23
110	-569	-23
111	-568	-22
112	-562	-22
113	-554	-22
114	-553	-22
115	-544	-21
116	-535	-21
117	-532	-21
118	-528	-21
119	-519	-21
120	-513	-20
121	-505	-20
122	-498	-20
123	-492	-19
124	-483	-19
125	-474	-19
126	-469	-18
127	-462	-18
128	-455	-18
129	-447	-18
130	-440	-17
131	-432	-17
132	-422	-17
133	-414	-16
134	-410	-16
135	-400	-16
136	-394	-15

137	-385	-15
138	-377	-15
139	-372	-14
140	-360	-14
141	-353	-14
142	-344	-13
143	-337	-13
144	-328	-13
145	-318	-12
146	-307	-12
147	-304	-12
148	-295	-11
149	-284	-11
150	-278	-11
151	-265	-10
152	-259	-10
153	-252	-10
154	-243	-9
155	-231	-9
156	-223	-9
157	-215	-8
158	-205	-8
159	-194	-7
160	-187	-7
161	-176	-7
162	-170	-6
163	-159	-6
164	-149	-6
165	-138	-5
166	-133	-5
167	-120	-4
168	-111	-4
169	-103	-4
170	-94	-3
171	-86	-3
172	-77	-3
173	-66	-2
174	-59	-2
175	-47	-2
176	-40	-1
177	-31	-1
178	-20	-0
179	-12	-0
180	-2	+0
181	+6	+1
182	+15	+1
183	+27	+1
184	+36	+2

185	+43	+2
186	+53	+3
187	+65	+3
188	+74	+3
189	+83	+4
190	+90	+4
191	+101	+4
192	+110	+5
193	+118	+5
194	+132	+6
195	+139	+6
196	+147	+6
197	+158	+7
198	+169	+7
199	+176	+7
200	+186	+8
201	+197	+8
202	+205	+9
203	+215	+9
204	+220	+9
205	+228	+10
206	+234	+10
207	+245	+10
208	+253	+11
209	+261	+11
210	+171	+12
211	+280	+12
212	+289	+12
213	+297	+13
214	+309	+13
215	+318	+13
216	+327	+14
217	+334	+14
218	+341	+14
219	+352	+15
220	+361	+15
221	+369	+15
222	+379	+16
223	+385	+16
224	+393	+16
225	+402	+17
226	+411	+17
227	+415	+17
228	+425	+18
229	+431	+18
230	+441	+18
231	+447	+19
232	+457	+19

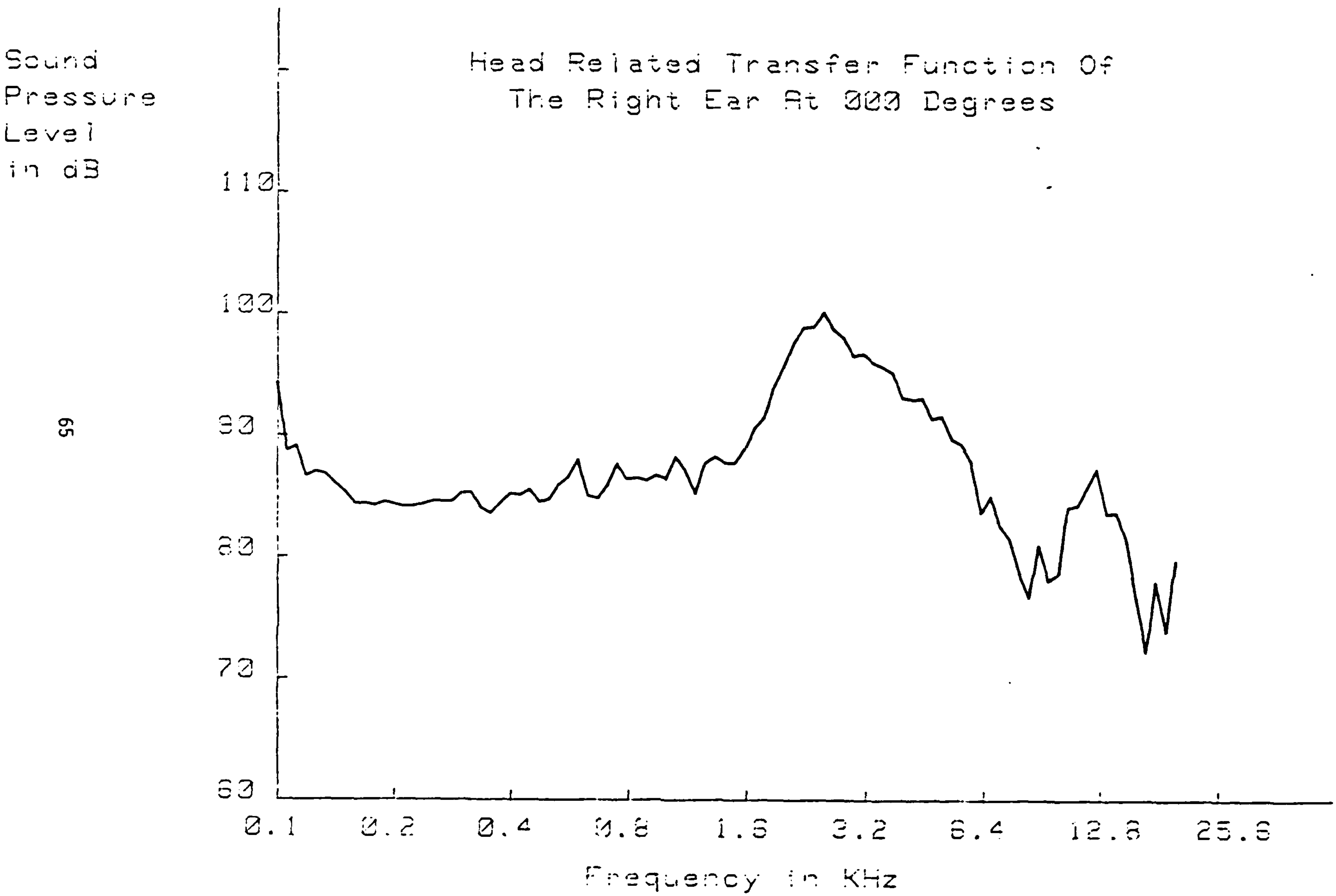
233	+464	+19
234	+469	+19
235	+477	+20
236	+485	+20
237	+495	+20
238	+497	+21
239	+510	+21
240	+514	+21
241	+519	+21
242	+524	+22
243	+535	+22
244	+541	+22
245	+549	+22
246	+553	+23
247	+559	+23
248	+564	+23
249	+571	+23
250	+578	+24
251	+581	+24
252	+590	+24
253	+596	+24
254	+599	+25
255	+607	+25
256	+614	+25
257	+619	+26
258	+628	+27
259	+645	+28
260	+677	+29
261	+712	+29
262	+718	+29
263	+719	+29
264	+724	+29
265	+721	+29
266	+722	+29
267	+717	+29
268	+720	+29
269	+714	+29
270	+715	+28
271	+711	+28
272	+708	+28
273	+704	+28
274	+700	+28
275	+697	+28
276	+691	+28
277	+688	+27
278	+682	+27
279	+681	+27
280	+673	+27

281	+664	+26
282	+661	+26
283	+653	+26
284	+648	+26
285	+640	+25
286	+632	+25
287	+624	+25
288	+618	+24
289	+610	+24
290	+599	+24
291	+592	+23
292	+581	+23
293	+574	+23
294	+565	+22
295	+557	+22
296	+550	+22
297	+541	+21
298	+534	+21
299	+528	+21
300	+518	+20
301	+511	+20
302	+5-1	+20
303	+492	+19
304	+486	+19
305	+477	+19
306	+471	+18
307	+462	+18
308	+455	+18
309	+445	+17
310	+437	+17
311	+427	+17
312	+420	+16
313	+410	+16
314	+404	+16
315	+396	+15
316	+384	+15
317	+375	+15
318	+367	+14
319	+358	+14
320	+348	+14
321	+341	+13
322	+333	+13
323	+323	+12
324	+311	+12
325	+3-3	+12
326	+295	+12
327	+289	+11
328	+278	+11

329	+271	+10
330	+262	+10
331	+255	+10
332	+245	+10
333	+238	+9
334	+229	+9
335	+222	+8
336	+211	+8
337	+205	+8
338	+192	+7
339	+193	+7
340	+178	+7
341	+170	+6
342	+162	+6
343	+151	+6
344	+144	+5
345	+133	+5
346	+129	+5
347	+118	+4
348	+108	+4
349	+101	+4
350	+93	+3
351	+83	+3
352	+76	+3
353	+68	+2
354	+59	+2
355	+43	+2
356	+35	+1
357	+27	+1
358	+17	+1
359	+9	+0
360	+1	+0

APPENDIX D

HEAD RELATED TRANSFER FUNCTIONS



Head Related Transfer Function Of
The Right Ear At 000 Degrees
Sound
Pressure
Level
in dB
110
100
90
80
70
60
0.1
0.2
0.4
0.8
1.6
3.2
6.4
12.8
25.6
Frequency in KHz

FIGURE D-1

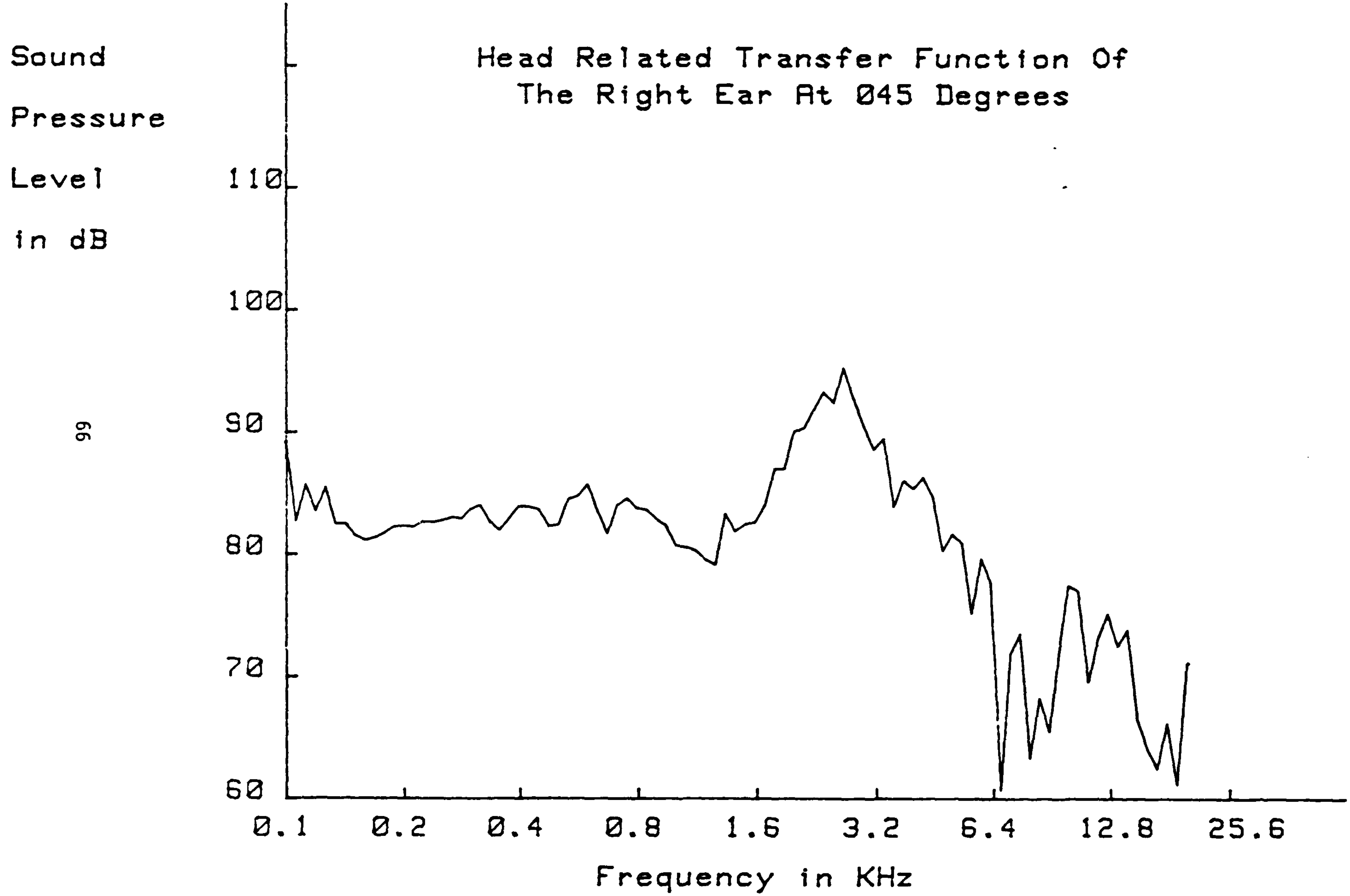
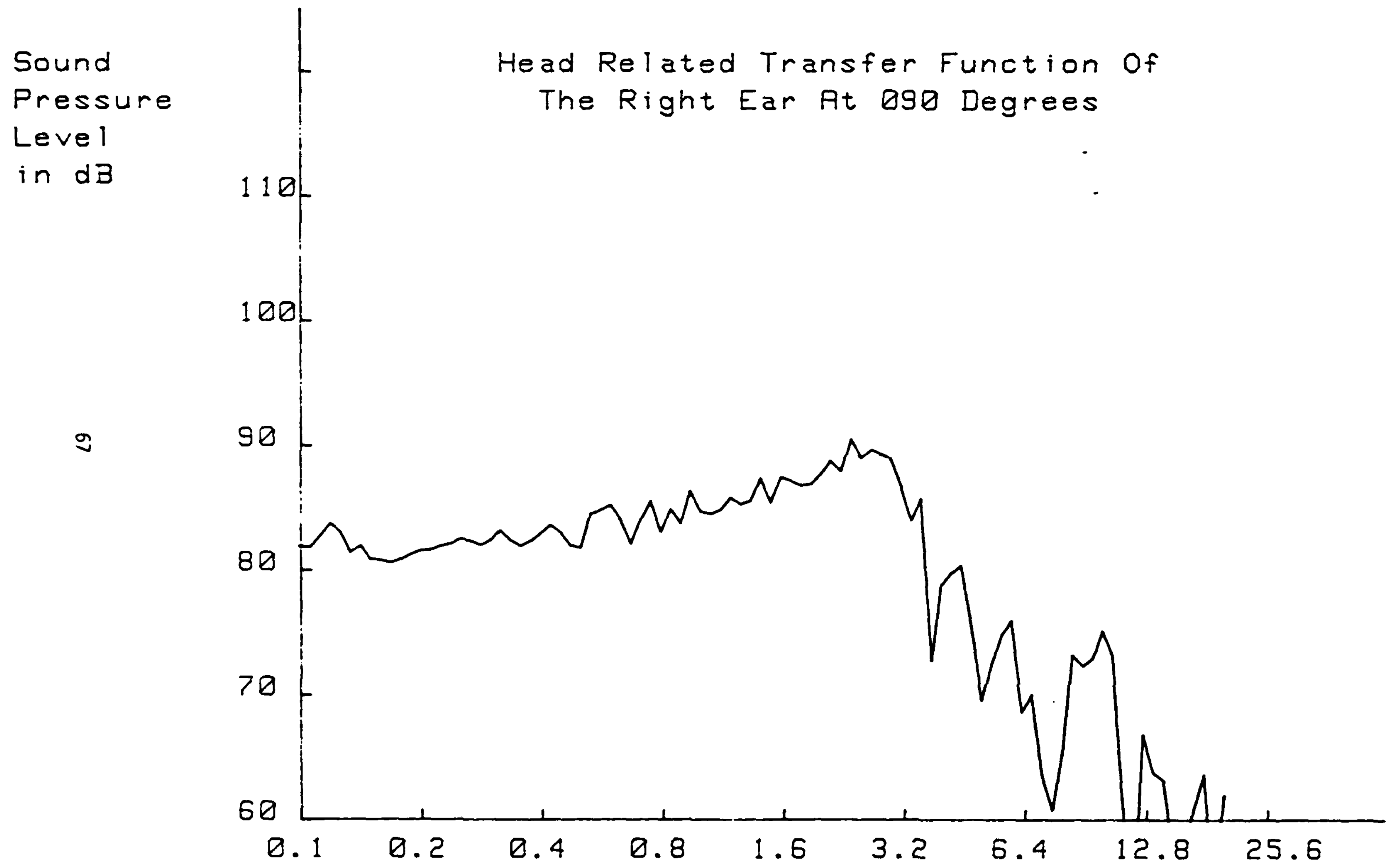
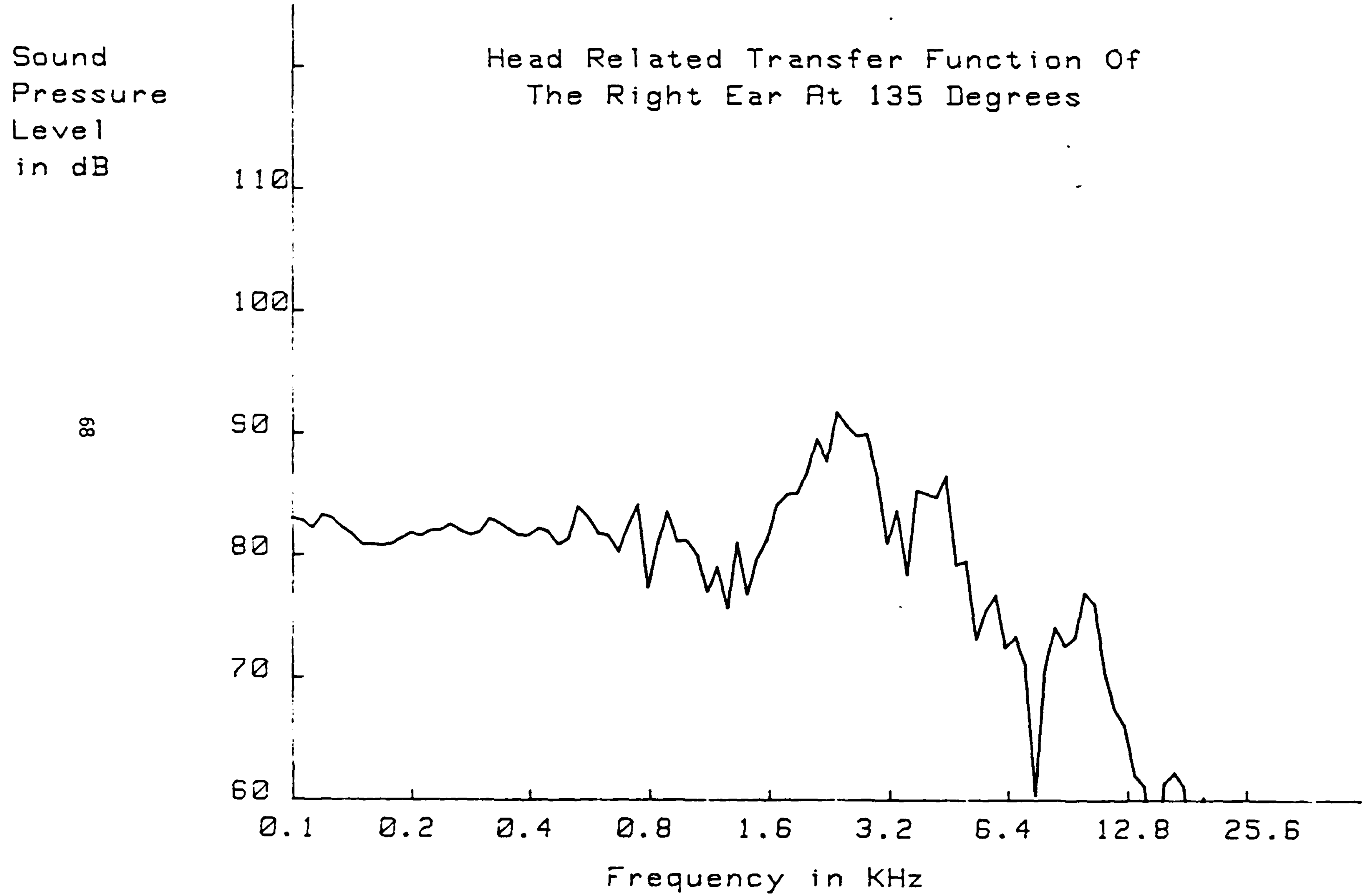

Head Related Transfer Function Of
The Right Ear At 045 Degrees
Sound
Pressure
Level
in dB
110
100
90
80
70
60
0.1
0.2
0.4
0.8
1.6
3.2
6.4
12.8
25.6
Frequency in KHz

FIGURE D-2



Head Related Transfer Function Of
The Right Ear At 090 Degrees
Sound
Pressure
Level
in dB
110
100
90
80
70
60
0.1
0.2
0.4
0.8
1.6
3.2
6.4
12.8
25.6
Frequency in KHz

FIGURE D-3


Head Related Transfer Function Of
The Right Ear At 135 Degrees
Sound
Pressure
Level
in dB
110
100
90
80
70
60
0.1
0.2
0.4
0.8
1.6
3.2
6.4
12.8
25.6
Frequency in KHz
FIGURE D-4

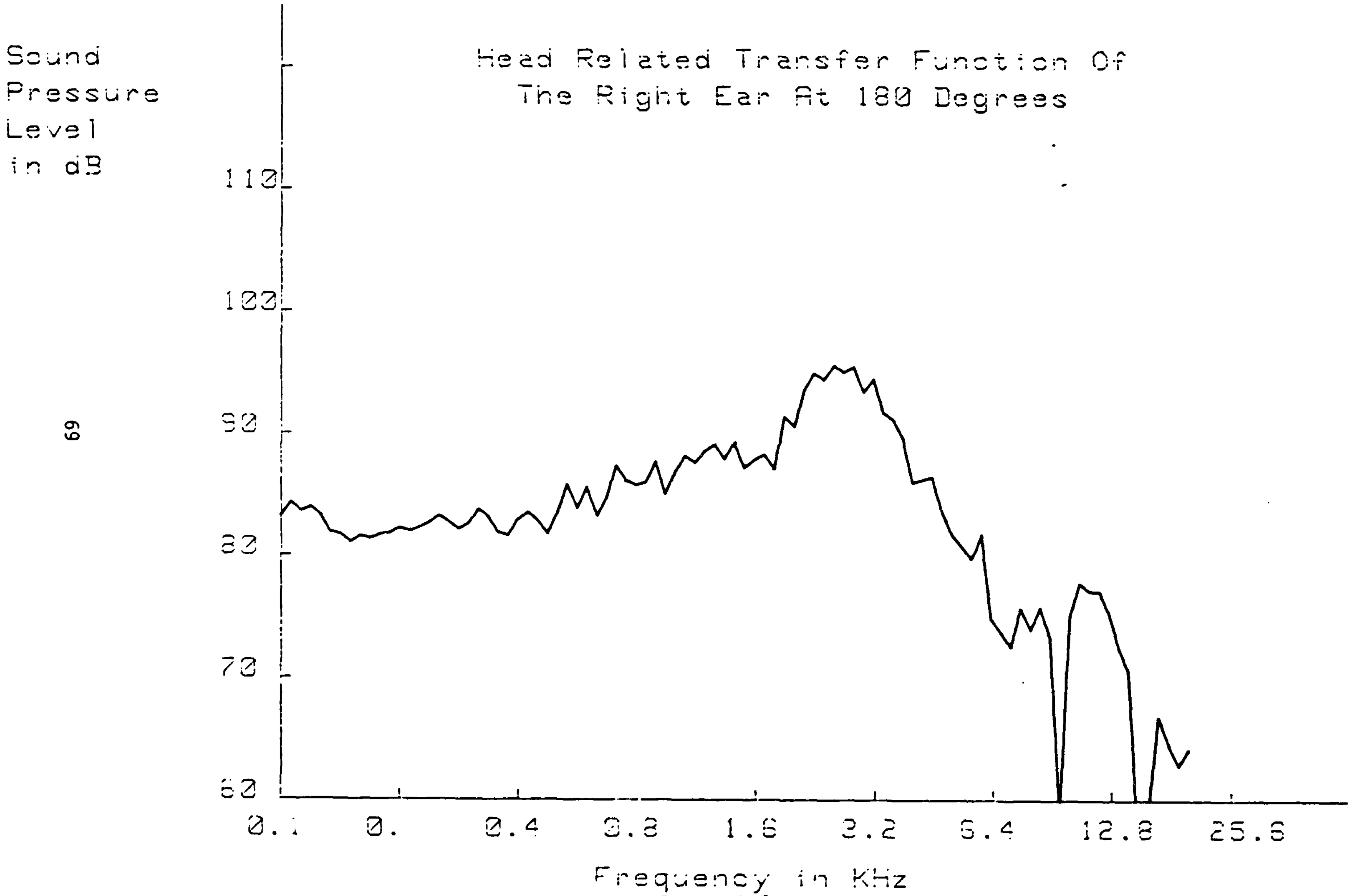
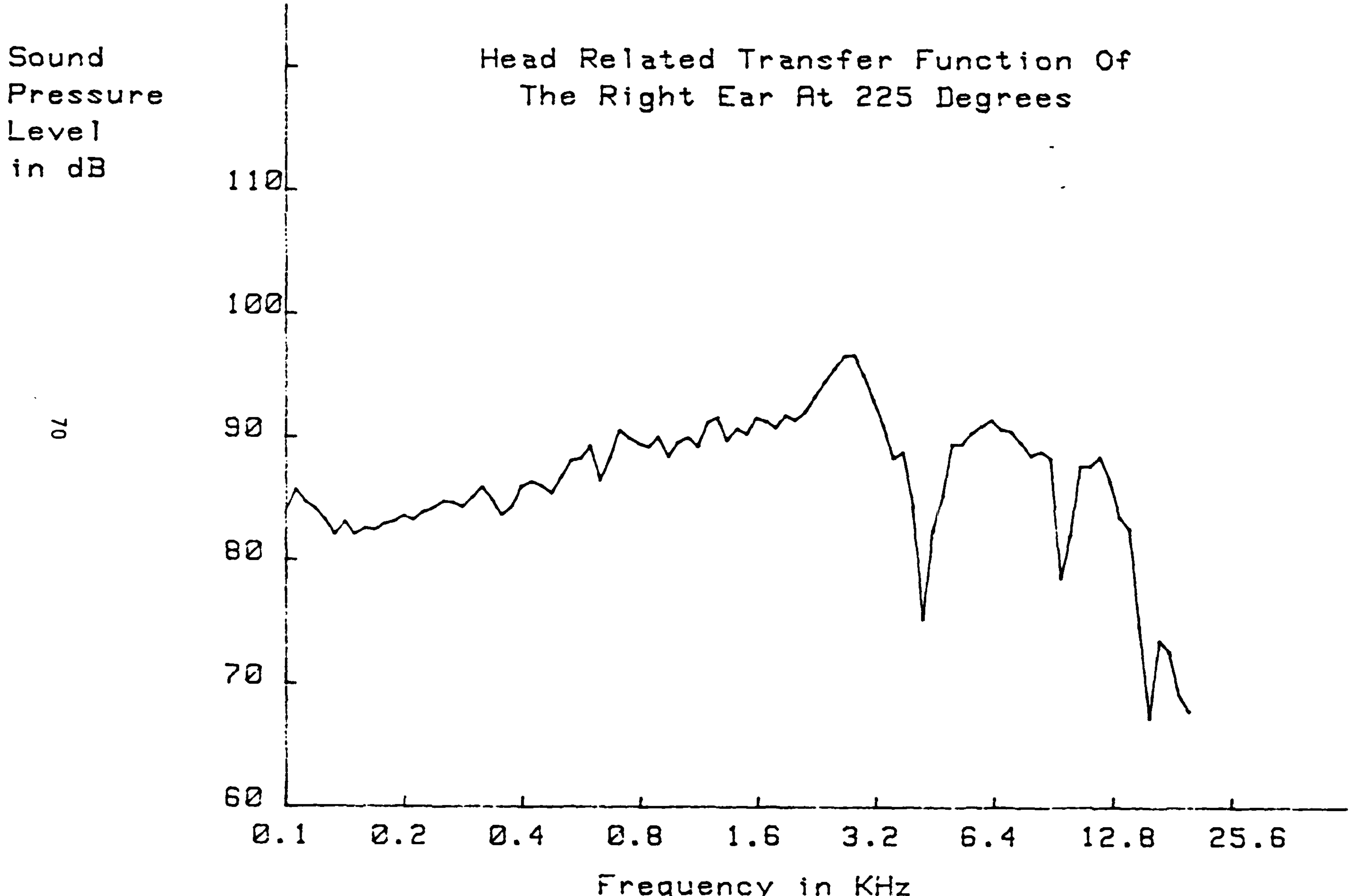

Head Related Transfer Function Of
The Right Ear At 180 Degrees
Sound
Pressure
Level
in dB
110
100
90
80
70
60
0.1
0.
0.4
3.8
1.6
3.2
6.4
12.8
25.6
Frequency in KHz

FIGURE D-5


Sound
Pressure
Level
in dB
Head Related Transfer Function Of
The Right Ear At 225 Degrees
110
100
90
80
70
60
0.1
0.2
0.4
0.8
1.6
3.2
6.4
12.8
25.6
Frequency in KHz
FIGURE D-6

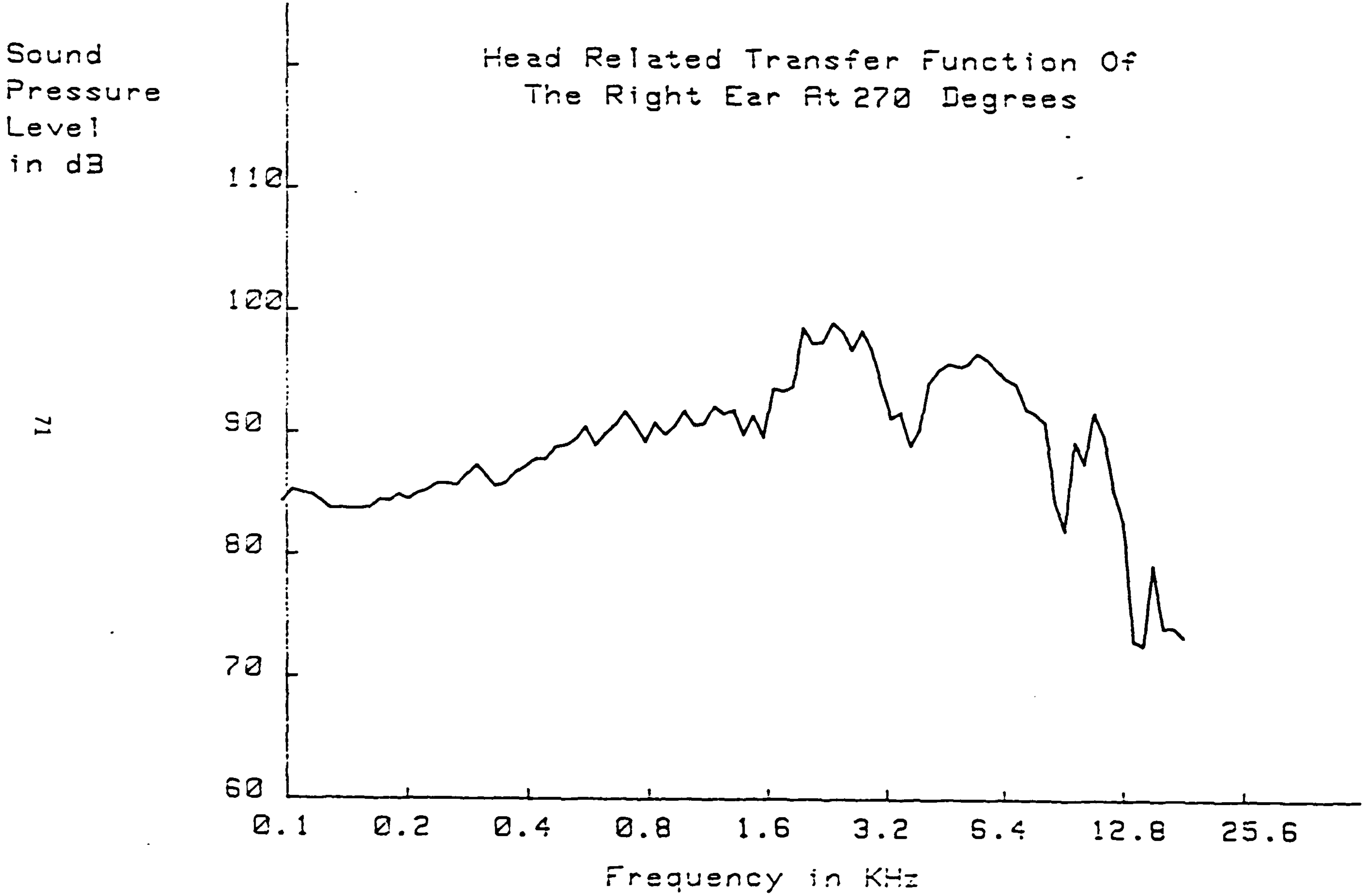
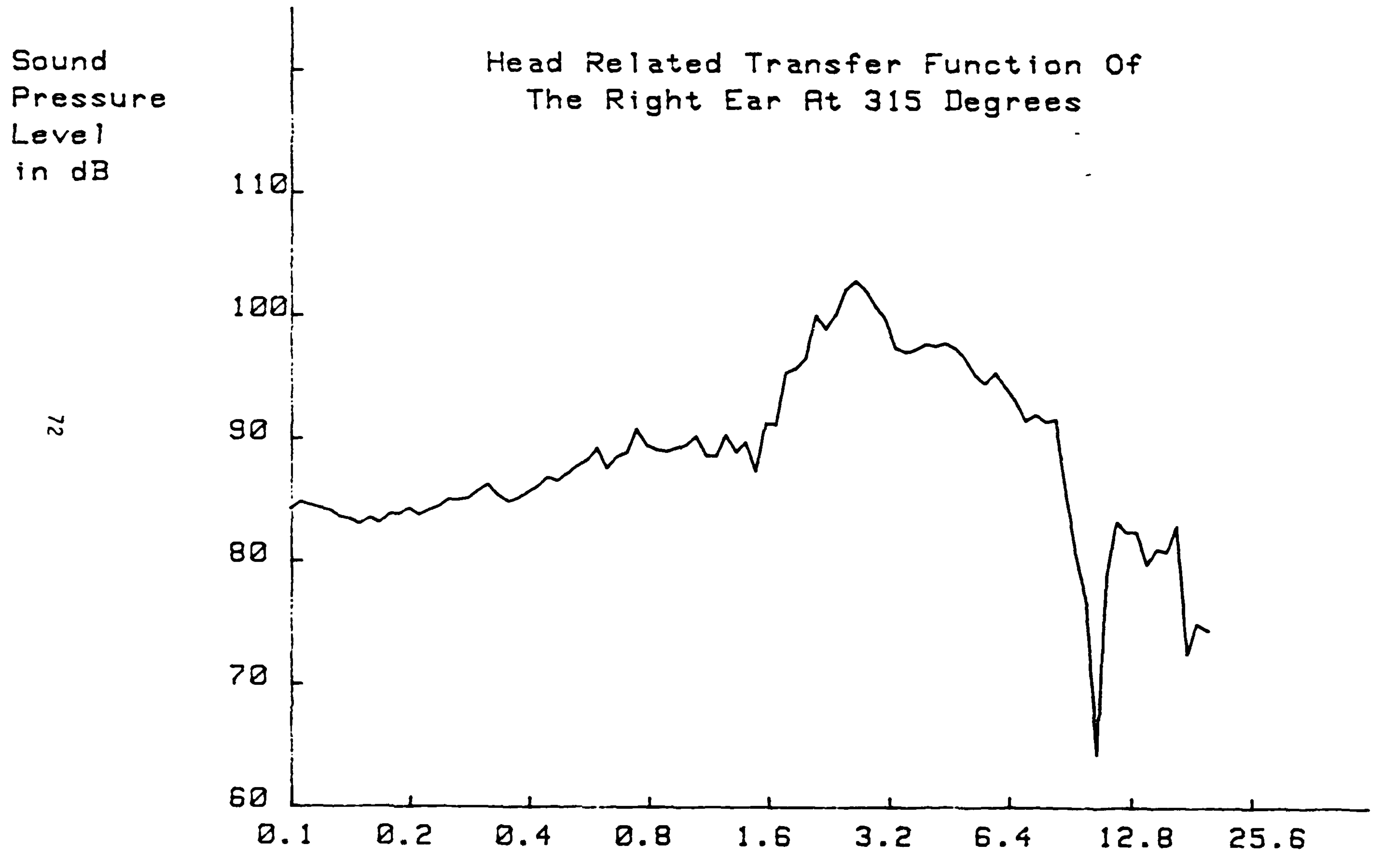

Head Related Transfer Function Of
The Right Ear At 270 Degrees
Sound
Pressure
Level
in dB
110
120
90
80
70
60
0.1
0.2
0.4
0.8
1.6
3.2
6.4
12.8
25.6
Frequency in KHz

FIGURE D-7



Head Related Transfer Function Of
The Right Ear At 315 Degrees
Sound
Pressure
Level
in dB
110
100
90
80
70
60
0.1
0.2
0.4
0.8
1.6
3.2
6.4
12.8
25.6
Frequency in KHz

FIGURE D-8

Bibliography

1. Batteau. D. W. "The role of the pinna in human localization," Proceedings of the Royal Society, London, Series B, 165, pg 158-180, London, England, 1967.

2. Batteau. D. W., Plante, R. L., Spencer, R. H., & Lyle, W. E. "Localization of sound: Part 3. A new theory of human audition," (Report No. TP3109, Part 3), China Lake, CA: U.S. Naval Ordnance Test Station, 1963.

3. Blauert, J. "Sound localization in the median plane," Acoustica, Vol 22, pg 205-213, 1969/1970.

4. Blauert, J. Spatial Hearing, The MIT Press, Cambridge, Massachusetts, 1983.

5. Burkhard, M. D., & Sachs, R. M. "Anthropometric manikin for acoustic research. Journal of the Acoustical Society of America, Vol 58, pg 214-222, 1975.

6. Colburn, H. S. "Binaural interaction and localization with various hearing impairments," Danevox Symposium, Copenhagen, Denmark, June 1982.

7. Deatherage, B. H., & Hirsh, I. J., "Auditory localization of clicks," Journal of the Acoustical Society of America, Vol 31, pg 486-492, April 1959.

8. Dirks, D. D. & Gilman, S. "Exploring azimuth effects with an anthropometric manikin," Journal of the Acoustical Society of America, Vol 66, pg 696-701, 1979.

9. Doll, T. J., Gerth, J. M., Engelman, W. R., Folds, D. J. Development of Simulated Directional Audio for Cockpit Applications, AAMRL-TR-86-014, January 1986.

10. Durlach, N. I. and Pang, X. D. "Interaural magnification," Journal of Acoustical Society of America, Vol 80, pg 1849-1850, December 1986.

11. Ericson, Mark A. and McKinley, R. L. Laboratory Data, Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB OH, October 1988.

12. Fechner, G. T. Elemente der Psychophysics [Elements of Psychophysics], Breitkopk und Hartel, Leipzig, 1860.

13. Flannery, R., & Butler, R. A. "Spectral cues provided by the pinna for monaural localization in the horizontal plane," Perception and Psychophysics, Vol 24, pg 438-444, 1981.

14. Freedman, S. J., & Fisher, H. G. "The role of the pinna in auditory localization," In S. J. Freedman (Ed.), The Neurophysiology of Spatially Oriented Behavior, pg 135-152, Homewood, IL: Dorsey Press, 1968.

15. Gatehouse, R. W. Localization of Sound: Theory and Applications, Amphora Press, Groton, Connecticut, July 1979.

16. Hebrank, J., & Wright, D. "Are two ears necessary for localization of sound sources on the median plane?," The Journal of the Acoustical Society of America, Vol 56, pg 935-938, 1974.

17. Hirsh, I. J. "The relation between localization and intelligibility," Journal of the Acoustical Society of America, Vol 22, pg 196-200, March 1950.

18. Kaiser, J. F. "Nonrecursive digital filter design using the I.-sinh window Function," Proceedings IEEE International Symposium on Circuits and Systems, April 1974.

19. Kuhn, G. R. "Model for the interaural time differences in the azimuthal plane," The Journal of the Acoustical Society of America, Vol 62, pg 157-167, 1977.

20. Lambert, R. M. "Dynamic theory of sound-source localization," The Journal of the Acoustical Society of America, Vol 56, pg 165-171, 1974.

21. Mehrgardt, S., & Mellert, V. "Transformation characteristics of the external human ear," The Journal of the Acoustical Society of America, Vol 61, pg 1567-1576, 1977.

22. Mills, A. W. "On the minimum audible angle," The Journal of the Acoustical Society of America, Vol 30, pg 237-246, 1958.

23. Morimoto, M., & Ando, Y. "On the simulation of sound localization," In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application, pg 85-98, Groton, CT: Amphora Press, 1982.

24. Musicant, A. D., & Butler, R. A. "The influence of pinnae-based spectral cues on sound localization," The Journal of the Acoustical Society of America, Vol 75, pg 1195-1200, 1984.

25. Musicant, A. D., & Butler, R. A. "The psychophysical basis of monaural localization," Hearing Research, Vol 14, pg 185-190, 1984.

26. Oppenheim, A. V., & Schafer, R. W. Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1975.

27. Perrott, D. R., & Musicant, A. D. "Dynamic minimum audible angle: Binaural spatial acuity with moving sound sources," The Journal of Auditory Research, Vol 21, pg 287-295, 1981.

28. Pollack, I., & Rose, M. "Effect of head movement on the localization of sounds in the equatorial plane," Perception and Psychophysics, Vol 2, pg 591-596, 1967.

29. Searle, C. L., Braida, L. D., Cuddy, D. R., & Davis, M. F. "Binaural pinna disparity: Another auditory localization cue," The Journal of the Acoustical Society of America, Vol 57, pg 448-455, 1975.

30. Schafer, R. W. Digital Filter Design Package, Version 2.12, Atlanta Signal Processors, Inc., Atlanta, GA, June 1987.

31. Shaw, E. A. G. "The external ear," in W. D. Keidel & W. D. Neff (Eds.), Handbook of Sensory Psychology, Vol 5, pg 455-490, Springer-Verlag, New York, 1974.

32. Shaw, E. A. G. "Transformation of sound pressure level from the far field to the eardrum in the horizontal plane," *The Journal of the Acoustical Society of America*, Vol *56*, pg 1848-1861, 1974.

33. Shelton, B. R., Rodger, J. C., & Searle, C. C. "The relation between vision, head motion and accuracy of free-field auditory localization," *The Journal of Auditory Research*, Vol *22*, pg 1-7, 1982.

34. Thurlow, W. R., Mangels, J. W., & Runge, D. S. "Head movements during sound localization," *The Journal of the Acoustical Society of America*, Vol *42*, pg 489-493, 1967.

35. Thurlow, W. R., & Runge, P. S. "Effect of induced head movements on localization of direction of sounds," *The Journal of the Acoustical Society of America*, Vol *42*, pg 480-488, 1967.

36. Wallach, H. "The role of head movements and vestibular and visual cues in sound localization," *Journal of Experimental Psychology*, Vol 27, pg 239-368, 1940.

37. Weiner, F. M. "On the diffraction of a progressive sound wave by the human head," *The Journal of the Acoustical Society of America*, Vol *19*, pg 143-146, 1947.

38. Wright, D., Hebrank, J. J. & Wilson, B. "Pinna reflections as cues for localization," The Journal of the Acoustical Society of America, Vol 56, pg 957-962, 1974.

VITA

Richard L. McKinley [redacted] He graduated from high school in Fairborn, Ohio in 1971 and attended Vanderbilt University, Nashville, Tennessee from which he received the Bachelor of Science degree in Biomedical Engineering in December 1975. He entered federal civil service May 1976 at the Aerospace Medical Research laboratory, Biodynamics and Bioengineering Division until entering the School of Engineering, Air Force Institute of Technology, in October 1983.

[redacted]

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distributed unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
AFIT/GE/ENG/88D-29

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
School of Engineering

6b. OFFICE SYMBOL *(If applicable)*
AFIT/ENG

7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS *(City, State, and ZIP Code)*
Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433

7b. ADDRESS *(City, State, and ZIP Code)*

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
Armstrong Aerospace Medical Research Laboratory

8b. OFFICE SYMBOL *(If applicable)*
AAMRL/BBA

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS *(City, State, and ZIP Code)*
AAMRL/BBA
Wright-Patterson AFB OH 45433-6573

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*
(U) CONCEPT AND DESIGN OF AN AUDITORY LOCALIZATION CUE SYNTHESIZER

12. PERSONAL AUTHOR(S)
Richard L. McKinley, B. S.

13a. TYPE OF REPORT
MS Thesis

13b. TIME COVERED
FROM ______ TO ______

14. DATE OF REPORT *(Year, Month, Day)*
1988 December

15. PAGE COUNT
87

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
23	02	
25	04	

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Auditory Localization; Bioacoustics; Digital Signal Processing; Communication; Human Factors

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

Title: CONCEPT AND DESIGN OF AN AUDITORY LOCALIZATION CUE SYNTHESIZER

Thesis Chairman: Dr. Matthew Kabrisky

10 Jan 89

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Dr. Matthew Kabrisky

22b. TELEPHONE *(Include Area Code)*
513-255-3576

22c. OFFICE SYMBOL
AFIT/ENG

DD Form 1473, JUN 86 *Previous editions are obsolete.*

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

This thesis describes the concept and design of an auditory localization cue synthesizer. The pertinent literature was reviewed and used to form the basis of the concept a to generate localization cues over headphones utilizing real-time solid state processor. The synthesizer accepts a single monaural input and processes the signal separately for independent presentation to the left and right ears. The synthesizer uses a 3-space head tracking device to maintain a stable acoustic image when the listener moves his head. The design is complete to present localized stimuli in azimuth. A concept is described for generating stimuli in the three dimensional case for azimuth, elevation and distance. Details of the hardware and software design are in the appendices.

Laboratory methodology are described for deriving the necessary parameters of the synthesizer. Experimental data collected separately from this thesis demonstrate that the concept and design are viable for the azimuth case. Localization errors with the synthesizer are compared with free field errors obtained with 10 subjects. The results show that localization accuracy is essentially equal for the two conditions. Recommendations are presented for further research and development.